Development of Genomic Resources and Core Collections of Germplasm for Tea Breeding

Development of Genomic Resources and Core Collections of Germplasm for Tea Breeding

A Dissertation Submitted to
the Graduate School of Life and Environment Sciences,
the University of Tsukuba,
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Agricultural Science
(Doctoral Program in Advanced Agricultural Technology and Sciences)

Fumiya Taniguchi

CONTENTS

ABBREVIATIONS

AFLP	amplified fragment length polymorphism
AMOVA	analysis of molecular variance
CAPS	cleaved amplified polymorphic sequence
CTAB	cetyltrimethylammonium bromide
CV	coefficient of variation
DFR	dihydroflavonol 4-reductase
EST	expressed sequence tag
GO	Gene Ontology
H_E	expected heterozygosity
H_O	observed heterozygosity
LG	linkage group
MAS	marker-assisted selection
NARO	National Agricultural Research Organization
NIAS	National Institute of Agrobiological Sciences
NIVTS	NARO Institute of Vegetable and Tea Science
PCOA	principal-coordinates analysis
PCR	polymerase chain reaction
PIC	polymorphic information content
RAPD	random amplified polymorphic DNA
SNP	single-nucleotide polymorphism
SSR	simple sequence repeat
STS	sequence-tagged sites
TAE	Tris-acetate-EDTA buffer
TAIL	thermal asymmetric interlaced (PCR)
TAIR	The Arabidopsis Infromation Resource

CHAPTER 1 General introduction

1.1 Tea

This section describes the features that make tea an important beverage crop, botanical aspects of the plant, and its medicinal benefits.

1.1.1 One of the world most popular beverages

Tea is one of the most popular beverages around the world. It is grown in more than 40 countries, mainly in Asia, the Middle East, Africa, and South America (FAOSTAT: http://www.fao.org/statistics/databases/en/).

Tea is made from leaves of *Camellia sinensis* (L.) O. Kuntze, and is classified into four major types: green, semi-"fermented", black, and post-fermented (Chen and Chen 2012). To produce green tea, fresh leaves are steamed or pan fried to inactivate polyphenol oxidase. To produce semi-fermented tea, leaves are allowed to wither for several hours, and the partially oxidized leaves are then pan fried. To produce black tea, withering is followed by rolling and "fermentation" (oxidization) at high temperature (about 30°C) and humidity. In the manufacture of black tea and semi-fermented tea, the process of "fermentation" is in fact oxidization by polyphenol oxidase. In contrast, in the production of post-fermented tea, after heating and rolling, the leaves are fermented by fungi or bacteria. In addition, in Southeast Asia, fermented and pickled tea leaves are eaten. This wide variety of tastes and aromas has fascinated humans for a long time because all tea comes from the leaves of the one species.

The annual worldwide production of tea in 2011 was 467 $\times 10^6$ t (FAOSTAT: http://www.fao.org/statistics/databases/en/), an increase of 56% from 299 $\times 10^6$ t in 2000. The status of tea has increased too, not only as a caffeine source, but also as a drink with health benefits. In Japan, in contrast, where tea has been consumed for centuries and consumption increased greatly in the second half of the 20th century, consumption peaked in the mid-1990s and has since declined gradually (Survey of Household Economy in Japan: http://www.stat.go.jp/english/index.htm). The decline is due to the greater availability of alternative drinks, such as coffee, juice, and mineral water, and the frequency of tea drinking has also decreased owing to changing lifestyles of Japanese consumers (Teramoto 2001).

1.1.2 Botanical and phytochemical aspects of tea

Tea is an evergreen woody plant. Its origin is unclear, but is considered to potentially include southwestern China, northern Vietnam, Laos, and Myanmar (Hashimoto 2001). Tea is a species of the genus *Camellia* in the family Theaceae. *Camellia sinensis* has two major varieties, *sinensis* and *assamica*. Variety *sinensis* is characterized by small leaves, a shrubby shape, and a low content of tannin (12%–19% by dry weight; Takeda 2004). It is distributed in East Asia, including China. In contrast, var. *assamica* has large leaves, a tree shape, and a high content of tannin (14%–23%). It is distributed from India to Southeast Asia.

The natural habitat of *C. sinensis* is subtropical, and the best temperature range is 20 to 30 °C. Cold resistance varies, and the temperature at which cold damage appears ranges from –3 to +15 °C. *Camellia sinensis* needs year-round rainfall of 1000 to 1700 mm a year (Chen and Chen 2012).

Usually, tea is harvested two to five times per year, depending on shoot regrowth. When a shoot tip is harvested, the lateral buds below start to grow, and become the next shoots to be harvested.

Camellia sinensis prefers acidic soils, with an optimum pH of 5.5 (Tsuji et al. 1994). Its preference for the form of nitrogen is notable: whereas many crops prefer nitrate, *C. sinensis* prefers ammonium (Morita et al. 2004, 2008). It is also resistant to aluminum, which is usually toxic to plants (Morita et al. 2004, 2008).

Since *C. sinensis* is an outcrosser and is self-incompatible (Tomo et al. 1956), most seedlings are assumed to derive from outcrossing.

Chemical compounds present in the shoots are critical to the product. Tea is one of the few major plant species that synthesize caffeine, along with coffee and cacao (Ashihara et al. 2008). A high content of polyphenols is an important characteristic. Catechins are the major polyphenols, with a content of 10% to 25% by dry weight (Takeda 2002). The catechin group of compounds includes six major forms: catechin, epicatechin, gallocatechin, epicatechin gallate, epigallocatechin, and epigallocatechin gallate (Chen and Chen 2012, Takeda 2002). In black tea processing, the catechins are enzymatically oxidized and oligomerised to theaflavin or thearubigin (Chen and Chen

2012). Amino acids, present at 1% to 6% by dry weight in new shoots, are important to tea quality. Theanine, an amino acid unique to *C. sinensis*, provides an *umami* (savory) taste, making it a particularly important compound. Theanine is especially high in *gyokuro* and *tencha* teas (Takeo 1992).

1.1.3 Medicinal benefits of tea

Tea has long been known to have medicinal effects. Recent studies have revealed antioxidant (Higdon and Frei 2003), anticancer (Kuroda and Hara 1999), and antibacterial and antiviral effects of catechins (Friedman 2007). A methylated catechin, epigallocatechin-3-*O*-(3-*O*-methyl) gallate, has anti-allergenic effects (Sano et al. 1999), and the Japanese cultivar 'Benifuki' has as much as 1% methylated catechin by dry weight (Maeda-Yamamoto et al. 2001). Theanine was shown to have a relaxing effect (Juneja et al. 1999).

1.2 History of tea use

The history of tea use dates back to at least 2737 B.C.E. (Yamanishi 1995). During the Tang dynasty (618–907 C.C.E.), tea had spread around China, becoming one of the most important items of trade, and had reached Tibet along the ancient Tea Horse Road, which supported the trade of tea for Tibetan horses (Chen 2004). The Dutch East India Company introduced tea into Europe around 1630 (Jacobs 2009). England introduced tea cultivation from China into India in 19th century. Then, in 1823, indigenous teas were found in Assam. Their morphological differences from Chinese teas led to the recognition of the two varieties, *assamica* and *sinensis*.

Vietnam is an old tea production area. Chinese history books indicate that tea was being grown there in the 7th century C.C.E. (Ngoc 2012). Northern Vietnam is very close to the origin of tea and is one of the oldest tea production areas (Ngoc 2012). This makes it an important area for exploring tea genetic resources.

Japan and Korea have grown tea for almost as long, since Buddhist monks introduced seeds from China in the 7th century C.C.E. (Tanaka 2012). Because the climate of most of the Korean Peninsula is too cold, tea cultivation is restricted to the southern part of the peninsula, and production is small (Jeong and Park 2012). In

contrast, tea production in Japan is extensive.

Major new production areas are Sri Lanka, Turkey, Iran, Africa, and South America. Sri Lanka, now the second largest producer of black tea in the world, has no indigenous tea. The first tea was planted in 1824, and commercial cultivation began in 1867, when a Scot, James Taylor, established a plantation (Gunasekare 2012). Most tea seeds planted in Sri Lanka were brought from India and China.

Central Asia and the Middle East started tea cultivation since the 19th century. More recently, African countries such as Kenya and South American countries such as Argentina have rapidly increased tea production (FAOSTAT: http://www.fao.org/statistics/databases/en/). These new producers grow teas from China and India.

The first record of the introduction of tea into Japan concerns two famous monks, Kukai and Saicho, who brought seeds from China in the 9th century C.C.E.. In the 13th century, a Buddhist priest, Eisai, who brought green tea from China to Japan, wrote a book called *Kissa-Youjou-Ki* to evangelize tea drinking in Japan. In the 16th century, Senno-Rikyu promoted the drinking of powdered green tea, *matcha*, and the tea ceremony became popular among the aristocracy and the samurai. In the 18th century, Sohen Nagatani established the process of steaming green tea, in which young leaves were harvested, steamed, kneaded, and then dried. This method was the basis for the manufacture of *sencha*. In the late 19th century, the Japanese government encouraged domestic tea production and promoted the export of tea, and thus production increased dramatically. Although production decreased during the two world wars, it recovered after World War II with the postwar reconstruction of Japan and rapid economic growth (Tanaka 2012).

In southwestern Japan, the existence of wild tea plants, called *yamacha* (“mountain tea”), led to the belief that Japan had indigenous teas. However, it is now thought that all *yamacha* originated as escapes from tea fields, and it is accepted that tea cultivation in Japan started with seeds brought from China by Buddhist monks. Most landraces in Japan are thought to have originated from those seeds. Since those seeds were limited, the genetic diversity of Japanese tea is low owing to the founder effect (Tanaka 2012).

1.3 Tea breeding

To identify the problems in tea breeding, it is important to review the history of tea breeding. This section describes the history of tea breeding, globally and in Japan.

1.3.1 History of global tea breeding

This section presents examples of tea breeding in several major producing countries.

China, the world's largest producer (FAOSTAT: http://www.fao.org/statistics/databases/en/) and the probable origin of tea, has the oldest history of tea breeding. The first authentic literature on tea, *Cha Ching* (*Tea Classics*), written by Lu Yu in 760 to 770 C.C.E., describes cultivars with different colors. Much later, in the 1780s, a tea farmer in Fujian province developed a method of vegetative propagation. Subsequently, two famous cultivars were bred: 'Tieguanyin' (oolong) in the 1780s and 'Fuding Dabaicha' (green) in 1857 (Yao and Chen 2012). In China, both public research organizations and private breeders have engaged in tea breeding. Until the 1980s, most cultivars were clones selected from seedlings derived from landrace tea fields. Since then, the percentage of registered clones created through controlled hybridization has increased notably, from 9.1% in 1987 to 22.2% in 2002 (Chen et al. 2007). Advances in genetic and molecular sciences have now allowed rapid progress in tea breeding in China (Yao and Chen 2012).

In India, a scientific approach to tea breeding was introduced in 1930 at the Tocklai Experimental Station of the Indian Tea Research Association (Das et al. 2012). Since there was no prior history of tea breeding and no knowledge about reproduction system of tea, work began with broad objectives:

(a) Production of pure lines by natural selfing.
(b) Breeding of clonal seed varieties superior to the commercial *jats* (seed populations) of tea.
(c) Selection of elite bushes and vegetative multiplication as clones.

In the 1950s and 1960s, the major target for selection in tea breeding was morphology. Breeding objectives have since focused also on yield, resistance to biotic and abiotic stresses, and drinking quality. Since most tea in India is grown for black tea,

breeding is focused only on black tea cultivars (Das et al. 2012).

In Sri Lanka, a scientific approach to tea breeding started with the establishment of the Tea Research Institute of Ceylon (now the Tea Research Institute of Sri Lanka) (Gunasekare 2012). In 1937, the Institute initiated the vegetative propagation of clonal cultivars and a planned breeding program. Until the 1950s, most tea cultivars were clones of selected seedlings derived from introduced seeds. In the 1980s, breeders began to use controlled hybridization.

In addition, Korea (Jeong and Park 2012), Vietnam (Ngoc 2012), Indonesia (Sriyadi et al. 2012), Bangladesh (Kahn 2012), Turkey (Ercisli 2012), Kenya (Kamunya 2012), and other countries have recently begun to breed cultivars by artificial crossing.

1.3.2 History of tea breeding in Japan

Tea breeding in Japan started with private breeders who selected superior individuals from seedling fields of landraces (Tanaka 2012). Organized breeding by public institutes started about a hundred years ago. Because tea became an important export item, the Japanese government supported the domestic production of tea and research on tea. Initially, cultivars were selected from separate populations of domestic landraces or of germplasms introduced from overseas. The cultivar ‘Yabukita’, which today has the largest share in Japan at about 76%, was selected from the field of a famous tea farmer, Hikosaburo Sugiyama in 1908, and propagated vegetatively. Planting of ‘Yabukita’ spread mainly since the 1970s. At the same time, the method of tea harvesting changed from hand-plucking to machine harvesting. This change accelerated the change from tea fields with seedling plants with unsynchronized sprouting to fields with clonal cultivars with synchronous bud break amenable to machine harvesting. ‘Yabukita’ was the best cultivar available at that time, and was therefore widely propagated. Although more than 60 clonal cultivars have since been released, they have not supplanted ‘Yabukita’.

1.3.3 Problems for the tea industry in Japan

In 2013, 42,800 ha of tea fields in Japan produced 82,800 tonnes of tea (Statistics on Crop, Statistics Department, Ministry of Agriculture, Forestry and

Fisheries, Japan: http://www.maff.go.jp/j/tokei/). Although tea production in Japan increased until the mid-1990s, it has since declined, in contrast to the continued increase in global production (Taniguchi 2013). The decline is due to the greater availability of alternative drinks, such as coffee, juice, and mineral water, and the frequency of tea drinking has also decreased owing to changing lifestyles of Japanese consumers. The Japanese tea industry is facing other problems too. To improve their financial status, farmers need to reduce production costs and the use of pesticides. In addition, the use of excessive amounts of nitrogen fertilizers to increase yields and improve tea quality has polluted rivers and lakes (Tanaka 2012), so this must also be reduced. Increases in damage caused by frost, freezing, drought, and typhoons are evident. And newly arrived invasive insects from overseas have created problems.

In spite of the decline in domestic production, exports continue to increase, as the global market plays an important role for the Japanese tea industry. However, when Japanese tea is exported to countries where agricultural chemicals specific to tea are not used, there is a risk that residues could exceed the limits set by the positive-list system, because the chemicals are not registered in the importing country. This situation creates a role for “organic” or pesticide-free tea, but because tea cultivars grown in Japan have poor resistance to pests and diseases, yield and quality suffer without the use of agricultural chemicals.

1.3.4 Problems of tea breeding in Japan

Today, ‘Yabukita’ accounts for 76% of the domestic tea production area (Tanaka 2012). It has contributed to the modernization of the Japanese tea industry, yet it has many problems. The most serious is that ‘Yabukita’ is sensitive to most pests and diseases (Taniguchi 2013), preventing efforts to reduce pesticide use. In addition, if farmers grow only one cultivar, the timing of harvest is concentrated in a short period, making it difficult to harvest the entire crop at the best timing. Although consumer preferences are diversifying, the dominance of this cultivar makes it difficult to make teas with different aromas and tastes. Thus, it is necessary to develop new tea cultivars with improved traits. However, despite the release of more than 60 new cultivars in Japan, the continued dominance of ‘Yabukita’ is due to the potential loss of production

and income for the 4 to 5 years it takes for new plants to grow to maturity.

1.3.5 Challenges to tea breeding in Japan

The solution to the current problems in tea breeding is to develop cultivars that are overwhelmingly superior to existing cultivars in one or more traits and to improve their economic advantage. Although this task requires us to improve several traits and is difficult to realize in a short time, continuous improvement of genetic potential is crucially important.

Essentially, breeding is the accumulation of useful genes from gene pools. Crop improvement occurs through selection operating on genetic variability. Selection can be intense and has resulted in major improvements in genetic variability. However, continuous success in plant breeding can be realized only insofar as a new variability is available for selection (Cooper et al. 2001). Without this variability, breeding decreases genetic diversity (Tanksley and McCough 1997). In general, low diversity in the gene pool causes improvement to plateau. Further, genetic uniformity renders plants vulnerable to disease or insect epidemics, such as the Irish potato famine of 1846 that was caused by late blight (Phytophthora infestans), and the decimation of the US maize crop in 1970 by southern leaf blight (Bipolaris maydis) as a result of cytoplasmic uniformity (Cooper et al. 2001).

In tea breeding in Japan, the introduction of genetic diversity has been limited because of the long generation time of tea. Parents have been limited to a few elite cultivars, and thus the genetic diversity of the breeding population is low. Therefore, few genes are available. In addition, since it takes a long time to grow tea plants until agronomic traits can be phenotyped, and large fields are needed, it is not easy to select superior individuals efficiently. Despite tea's status as one of the most popular drinks in the world and its long history, tea breeding has not yet achieved the results obtained for other major crops.

To improve tea breeding continuously, therefore, it is necessary to devise a breeding system that efficiently accumulates many useful genes from various sources.

1.4 Technical problems to solve

To accelerate the introduction and accumulation of diverse useful genes in tea breeding, the following tasks are necessary:

1) Efficient screening of breeding materials from germplasm collections.
2) Identification of useful genes by genetic analysis.
3) Efficient accumulation of useful genes by DNA-marker-assisted selection and rapid generation advancement.

This section describes the current knowledge of genetic resources and genomic information of tea, and points out the technical problems to be solved in this field of study.

1.4.1 Genetic resources of tea

In Japan, parents used in tea breeding are strongly biased toward green tea cultivars derived from domestic landraces with low genetic diversity. Thus, only limited genes and combinations of genes have been used in tea breeding, and the results of breeding are inevitably limited. Germplasms from various origins should be used more.

Germplasm collections of tea are conserved in several countries, notably China, India, and Japan. In Japan, NARO Institute of Vegetable and Tea Science (NIVTS) holds a collection within the National Institute of Agrobiological Sciences (NIAS) Genebank. This collection, one of the largest germplasm collections of tea, was started about a hundred years ago, when tea, especially black tea, was an important export item for Japan. Its major objective was to introduce tea seeds from China and India for improvement of black tea quality.

However, in breeding for green teas, germplasms introduced from overseas have not been used fully. To enrich the genetic diversity of breeding populations and achieve various breeding objectives, it is essential to use more diverse genetic resources.

Although extensive collections are a good resource, it can be difficult to include many accessions at the same time in one experiment. For example, since most assays for pest resistance are laborious, breeders can screen only a limited number of accessions. Therefore, a key point is the selection of materials for screening from the whole collection. This can be facilitated by detailed information about the genetic

diversity of the resources. DNA markers make powerful tools for the evaluation of genetic diversity. In particular, SSR markers are used in many crops because they are highly polymorphic, multi-allelic, and stable (Kalia et al. 2011).

Core collections allow many germplasms to be managed efficiently. A core collection, as proposed by Frankel (1984), is a small subset which covers the genetic diversity of a full collection with a limited number of accessions. Core collections of many crops have been selected by using genotype data for DNA marker loci and used in breeding and research on genetic diversity (Kojima et al. 2005, Kaga et al. 2012, Balfourier et al. 2007, Díez et al. 2012, Le Cunff et al. 2008). Several studies of genetic diversity analyzed by using DNA markers in tea have been reported (Wachira et al. 2001; Kaundun and Matsumoto 2002, 2003; Yao et al. 2008, 2012; Ohsako et al. 2008; Raina et al. 2012; Fang et al. 2012), revealing high diversity in Chinese and Indian teas and low diversity in Japanese teas (Kaundun and Matsumoto 2003, Yao et al. 2012, Raina et al. 2012). However, previous studies were disadvantaged by using too few germplasms or a too-narrow range of origins. To date, no core collections from worldwide genetic resources of tea have been selected by using DNA markers.

1.4.2 Genome research in tea

As mentioned above, the key to improving tea breeding is to efficiently introduce an abundance of useful genes from diverse genetic resources into breeding programs. The genetic analysis of important agronomic traits needs support from DNA markers and linkage maps. The analysis of genetic diversity and selection of core collections also need a great number of DNA markers.

Recent progress in genome research has enabled the genome sequencing of many organisms. This has made it possible to develop DNA markers for these organisms. It is also possible to develop DNA markers when a genome sequence is not available if that of a closely related species is available. However, there is no model organism or major crop closely related to *C. sinensis* for which genome sequencing has been completed. Furthermore, the genome of *C. sinensis* is very large, at 4 Gbp (Tanaka et al. 2006), and it is thought to contain a lot of repetitive sequences. Therefore, it is not feasible to develop DNA markers from the genome sequence of tea yet.

Although modern *C. sinensis* is diploid ($2n = 30$), it is thought to be paleopolyploid (Shi et al. 2010), which would allow plenty of repetitive sequences to have accumulated. The self-incompatibility and high heterozygosity of *C. sinensis* would make it easy to detect polymorphisms for DNA-marker alleles.

Researchers have reported the development of DNA markers, phylogenetic analyses, and the construction of linkage maps for tea. Ota and Tanaka (1999) constructed linkage maps using RAPD markers and pseudo-testcross theory. RAPD markers have also been used to investigate the genetic diversity of tea genetic resources (Tanaka et al. 1995, Kaundun and Matsumoto 2002 Wachira et al. 2001). AFLP and CAPS markers have been developed and used for analyses of genetic diversity (Raina et al. 2012, Kaundun and Matsumoto 2003). Although several linkage maps for the parents of mapping populations have been constructed by using pseudo-testcross theory with dominant markers such as RAPD markers (Tanaka et al. 1995, Ota and Tanaka 1999, Hackett et al. 2000), they have several weaknesses. For example, the number of linkage groups was not 15, the haploid chromosome number, the markers did not cover certain portions of the genome, and the relationship among the several linkage maps was not clear. To perform linkage analysis for the development of DNA markers for various agronomic traits, it is important to construct a reference linkage map that covers the entire genome and is compatible with various mapping populations. For this purpose, DNA markers that are highly polymorphic and applicable to various mapping populations are essential. SSR markers are suitable for the construction of such a map because they are codominant and highly polymorphic. Since mapping populations of *C. sinensis* are derived from crosses between heterozygous parents, a locus has up to four alleles. Multiallelic SSR markers can distinguish the four alleles. A reference map based on SSR markers is also useful for the analysis of genetic diversity, because the multiallelicity of SSRs is an excellent feature for genotyping various germplasms. In addition, if the markers for genotyping of germplasms are selected evenly from the entire genome, it is possible to investigate the genetic diversity of germplasms evenly across the genome.

It will take enormous cost and labor to sequence the entire genome of *C. sinensis* because it is very large. On the other hand, it is realistic to investigate only expressed genes. Once information about expressed genes is accumulated, it will

become possible to develop DNA markers using the sequence information and map them onto linkage maps. As expressed genes contain SSR motifs in their sequences, the information is a good resource for the development of SSR markers. Furthermore, functional annotation makes it possible to map the target genes by using candidate genes that are thought to be involved in the traits, in the “candidate gene approach” (Pflieger et al. 2001).

1.4.3 Objective of this dissertation

The first aim of this study was to develop genome resources to capture agronomically important genes and develop cultivars in which the genes are accumulated. The second aim was to select core collections from the worldwide tea germplasm collection to efficiently find germplasms which have useful genes.

By performing this study, I aimed at contributing to construct a modern tea breeding system that allows continuous and surefire improvement in performance of tea cultivars.

CHAPTER 2 Development of SSR markers from ESTs and an SSR-enriched genomic library in tea

2.1 Introduction

For the genetic analysis of important agronomic traits and analysis of the genetic diversity of germplasms, the development of DNA markers is a prerequisite. Such DNA markers must be highly polymorphic, detect genotypes stably, and be applicable to high-throughput analysis. SSR markers are co-dominant, highly poly3morphic, and stable. High-throughput analysis is possible with a DNA sequencer and software for fragment analysis. SSRs are therefore very suitable for both goals. SSR markers come in two types: genic, based on the sequences of expressed genes, and genomic, based on the sequences of SSR-enriched genomic libraries.

To date, several studies have developed SSR markers in tea. Before this study, about 130 genic SSRs (Jin et al. 2006, Sharma et al. 2009, Zhou et al. 2011) and only 50 genomic SSRs (Chen et al. 2000, Freeman et al. 2004, Hung et al. 2008, Yang et al. 2009) were available. Because the development of linkage maps for the genetic analysis of important agronomic traits requires hundreds of genome-wide markers, this is not enough for tea. Therefore, the present study concerned the large-scale development of genic and genomic SSRs. Both types were developed because the distribution of one or the other type could be limited in some regions of the genome.

To develop genic SSRs, it is necessary to sequence many expressed genes. EST analysis reads many partial sequences of clones in cDNA libraries. Although the large genome of *C. sinensis* hampers full-genome sequencing, it is feasible to analyze expressed genes, which number in the tens of thousands at most. So an efficient first step for the analysis of large-genome species such as tea is to survey the expressed genes.

Several EST analyses of tea have been reported. Chen L. et al. (2005) reported 1684 ESTs generated from tender shoots. Park et al. (2004) reported 588 ESTs isolated by suppression subtractive hybridization. Sharma and Kumar (2005) reported three drought-responsive ESTs obtained by differential display. Shi et al. (2011) reported

details of the transcriptome of *C. sinensis* that were generated by RNA-seq analysis using a high-throughput Illumina GA IIx sequencer. The ESTs reported in the first three studies were derived from green tissues only. The RNA-seq data reported by Shi et al. (2011) were generated from seven different organs, including young roots, flower buds, and immature seeds, but the RNAs were mixed before analysis, and thus the origin of each transcript could not be identified.

The study described in this chapter reports the development of 17,458 ESTs derived from seven cDNA libraries of young shoots, mature leaves, and roots of tea plants. To facilitate gene identification and functional studies, I performed Gene Ontology (GO; Ashburner et al. 2000) annotation of tea unigenes. EST-SSR markers developed from the EST data are highly polymorphic and thus transferable to many *Camellia* species. In addition, I developed genomic SSRs from an SSR-enriched genomic library and surveyed their polymorphisms.

2.2 Materials and methods

2.2.1 Plant materials

Organs for RNA isolation were collected from tea plants growing at the Makurazaki Tea Research Station, NIVTS, Kagoshima, Japan. Young roots came from 15-d-old seedlings derived from natural crosses of *C. sinensis* cv. 'Sayamakaori'. Tap roots and lateral roots were harvested from 30-d-old seedlings. Young leaves, terminal buds , and young stems of growing shoots with two leaves and a bud were harvested from field-grown 'Sayamakaori' in April of the first flush (first harvest) season, along with mature leaves that developed the previous year. Sixteen accessions of *C. sinensis* (Table 2.1) and 14 other *Camellia* species (Table 2.2) were used for EST-SSR analysis.

2.2.2 Preparation of total RNA and cDNA library construction

Total RNAs from aboveground tissues were extracted using Trizol Reagent (Life Technologies, USA). Total RNAs from root tissues were extracted using an RNeasy Plant Mini Kit (Qiagen, Germany).

For the construction of a cDNA library from the young-roots RNA, total RNA was dephosphorylated and decapped with a GeneRacer kit (Life Technologies). It was

then ligated with GeneRacer RNA Oligo and reverse-transcribed with SuperScript II reverse transcriptase (Life Technologies). After first-strand cDNA synthesis, the RNA was degraded with RNase H. cDNA was amplified by PCR with the forward (5′-CGACTGGAGCACGAGGACACTGA-3′) and the reverse (5′-GCTGTCAACGATACGCTACGTAACG-3′) primers at an initial 94 °C for 2 min; 20 cycles at 94 °C for 20 s, 56 °C for 30 s, and 72 °C for 10 min; and a final extension at 72 °C for 10 min. To enrich the content of long cDNAs, the PCR products were separated by agarose gel electrophoresis, and products longer than 1000 bp were isolated and cloned into the pGEM-T Easy vector (Life Technologies), which was then transformed into *Escherichia coli* strain DH5α cells.

For the construction of cDNA libraries from the other organs, double-stranded cDNA was synthesized with a SMART cDNA Library Construction Kit (Clontech, USA), digested with restriction enzyme *Sfi*I, and size-fractionated in a CHROMA-SPIN 400 column (Clontech). The cDNA fragments were directionally ligated into an *Sfi*I-digested pTriplEx2 vector. The ligation mixture was electroporated into *E. coli* DH5α competent cells.

2.2.3 DNA sequencing

Both ends of cDNAs from the YR library were sequenced using primer T7 (5′-TAATACGACTCACTATAGGG-3′) or SP6 (5′-ATTTAGGTGACACTATAGAA-3′), and the 5′ ends of cDNAs from the YL, TB, YS, ML, TR, and LR libraries were sequenced using the 5′ λTriplEx2 sequencing primer (5′-TCCGAGATCTGGACGAGC-3′). Cycle sequencing reactions were performed using a BigDye Terminator Cycle Sequencing Kit (Life Technologies), and capillary electrophoresis used an ABI 3730*xl* or 3130*xl* sequencer (Life Technologies).

2.2.4 Sequence analysis

Base-calling of sequence reads was performed using the KB Basecaller program (Life Technologies). Ambiguous sequences were removed using the Sequencing Analysis program (Life Technologies), and vector sequences were trimmed using the Cross_match program (http://www.phrap.org/). Sequences of <100 bp were

then eliminated from the analysis. A total of 17,458 ESTs were generated and submitted to the DDBJ database (accession numbers AB361047 to AB361052, AB461364 to AB461372, AB485966-AB485975, AB505865-AB505873, and FS943336 to FS960759). The 17,458 ESTs were assembled using the phrap program (http://www.phrap.org/). If the 5′ and 3′ reads derived from the same clone in the young-roots library belonged to different contigs, or both reads were singletons, or one read was a member of a contig and the other was a singleton, then the contigs or singletons were treated as a single scaffold.

The nucleotide sequences of the unigenes were compared using the BLASTX algorithm (Gish and States 1993) against the non-redundant protein sequences in GenBank (http://www.ncbi.nlm.nih.gov/genbank/), the UniProt database (http://www.uniprot.org/), the Arabidopsis proteome database (TAIR8; http://www.arabidopsis.org/), and amino acid sequences deduced from the rice genome sequence (IRGSP/RAP build 5; http://rapdb.dna.affrc.go.jp/download/). Unigenes were functionally annotated using the Blast2GO program (Conesa et al. 2005). GO Slim annotations of unigenes were also generated with Blast2GO using the plant GO Slim mapping program provided by TAIR (http://www.arabidopsis.org/).

2.2.5 Digital analysis of expression

I selected 144 unigenes that were generated by the assembly of 10 or more independent ESTs and used them for expression profiling based on the number of ESTs within each library. Differential expression levels were tested with the Audic and Claverie statistical test in the IDEG6 software (Romualdi et al. 2003). To eliminate false positives, I used Bonferroni's correction for the adjustment of multiple comparisons. Sixty-seven unigenes that were expressed differently among the seven libraries were clustered using Hierarchical Clustering Explorer v. 3.0 software (http://www.cs.umd.edu/hcil/hce/).

2.2.6 Identification of EST-SSRs

Using the tea unigene set as a target, I identified microsatellites with ≥6 repeats, each repeat unit being ≥3 repeats of dinucleotides or trinucleotides, using the

Read2Marker program (Fukuoka et al. 2005). I also designed PCR primers for amplification of EST-SSRs using Read2Marker. EST-SSR markers were named with 'MSE'.

2.2.7 Identification of genomic SSRs

Genomic SSR markers were developed from microsatellite-enriched libraries. These libraries, enriched for GA and GT, were developed by Genetic Identification Services Inc. (Chatsworth, CA, USA) from DNA of a line KM62, extracted by using a modified CTAB method (Tanaka et al. 2001). Markers were sequenced on a model 310, 3130*xl*, or 3700 Genetic Analyzer (Applied Biosystems). The primer sets were designed with the Primer Express (Applied Biosystems) or Read2Marker software. Genomic SSR markers were named with 'MSG' or 'TMS-LA'.

2.2.8 Analysis of SSRs

PCR was performed in a 10-µl reaction mix including 20 ng of genomic DNA, 10× PCR Gold buffer (Life Technologies), 0.8 µL of 8 mM dNTPs, 0.1 U of AmpliTaq Gold Polymerase, 0.8 µL of 25 mM $MgCl_2$, and 1 µM of the forward and reverse primers. The PCR reactions were carried out in a GeneAmp 9600 thermal cycler (Life Technologies) according to the following touchdown PCR cycling program: 95 °C for 5 min; 95 °C for 1 min, 62 °C for 30 s, and 72 °C for 1 min; 13 cycles at decreasing annealing temperatures in decrements of 0.5 °C per cycle; 25 cycles of 95 °C for 1 min, 55 °C for 30 s and 72 °C for 1 min; and a final extension at 72 °C for 10 min. PCR products were directly labeled with fluorescence-labeled R110-ddUTP by the single-tube method (Inazuka et al. 1996). The labeled PCR products were analyzed on an ABI Prism 3130*xl* Genetic Analyzer, and the resulting allele data were analyzed with GeneMapper v. 3.7 software (Life Technologies). Polymorphism information content and heterozygosity information were calculated in the PowerMarker software (Liu and Muse 2005).

2.3 Results

2.3.1 Sequencing and assembly

Seven cDNA libraries were constructed from tea plant organs (Table 2.3). From the young roots cDNA library, 3072 clones were randomly selected and single-pass-sequenced from both ends. From each of the other six libraries, 2880 clones were sequenced from their 5′ ends. After removal of low-quality sequences and vector trimming, the resulting data set contained 17,458 sequences with an average length of 481 bp (Table 2.4). The GC content of the 17,458 ESTs (8,391,523 bases) was 44.0%. Assembly of these 17,458 high-quality ESTs into contigs in phrap resulted in 2227 contigs and 3477 singletons. Some 5′ and 3′ reads from the same clones from the young roots library were not assembled into the same contigs; in such cases, the contigs and singletons that contained such reads were treated as scaffolds. As a result, 442 scaffolds, 1851 contigs, and 2969 singletons were generated. Together, the 5262 sequences were used for further analysis as a 5.3-k tea unigene set. Among these sequences, 3372 unigenes (64.1%) were longer than 500 bp (Table 2.4). The assembly of ESTs in each cDNA library generated 846 to 1587 unique transcripts per library (Table 2.3).

On 5 August 2011, the NCBI GenBank database contained 14,246 ESTs and 34.5 $\times 10^6$ RNA-seqs from tea. Similarity searches of the 5.3-k unigene set were performed against the 14,246 ESTs and the 76,159 assembled sequences from the RNA-seq analysis of Shi et al. (2011), which had been deposited in the Transcriptome Shotgun Assembly Sequence Database at NCBI with accession numbers HP701085–HP777243. BLASTN searches of the 5.3-k unigene set with a cutoff value of 1e–10 identified 3340 unigenes (63.5%) with no matches among the 14,246 tea ESTs in GenBank, 1118 (21.2%) with no matches among the 76,159 assembled sequences of RNA-seqs, and 732 (13.9%) with no significant matches within either data set.

2.3.2 Similarity search and functional annotation

A BLASTX search against the GenBank non-redundant database (cutoff of $\leq 1e^{-6}$) returned significant hits for 3055 unigenes in the 5.3-k set (58.1%). Of those, 762 (24.9%) were annotated as hypothetical, predicted, putative, unknown, or unnamed proteins. BLASTX searches against the UniProt database and the complete protein sets of *Arabidopsis thaliana* and *Oryza sativa* (cutoff of $\leq 1e^{-6}$) found that 2484 (47.2%) of the 5.3-k unigene set encoded peptides with significant similarity to those in the UniProt

database, 3417 (64.9%) similar to *Arabidopsis* proteins, and 3673 (69.8%) similar to rice proteins.

GO annotation by Blast2GO annotated 2639 unigenes with 11,260 annotations, distributed among the main GO categories of Biological Process (4582), Molecular Function (3509), and Cellular Component (3169) (Fig. 2.1). Detailed annotation information is available as the Supplemental Table 1 at the following site, https://www.jstage.jst.go.jp/article/jsbbs/62/2/62_186/_article/supplement. There were 1191 unigenes annotated in all three GO categories.

To evaluate the usefulness of the 5.3-k unigene set as a gene resource for tea, I searched for unigenes involved in important horticultural and biological processes of tea, namely nitrogen (N) assimilation and amino acid metabolism, catechin and caffeine biosynthesis, photoresponse, and aluminum (Al) response (Table 2.5). In N assimilation, I found unigenes involved in primary assimilation of inorganic N, such as nitrate transporter, ammonium transporter, and glutamate synthetase annotations, and in amino acid metabolism. In catechin biosynthesis, I found unigenes encoding 10 enzymes, including phenylalanine ammonia-lyase and leucoanthocyanidin reductase. In addition, I identified unigenes encoding caffeine synthetase and some involved in Al response and photoresponse.

2.3.3 Digital analysis of gene expression

To reveal patterns of gene expression and correlations of expression patterns among organs, I analyzed the EST data using R statistics of the IDEG6 web tool (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) to identify unigenes that were differentially expressed. From the 5.3-k unigene set, 144 unigenes that consisted of >10 EST sequences were selected for analysis; of these, 67 showed significant differences in their expression profile among the libraries (Table 2.6). Cluster analysis using the Hierarchical Clustering Explorer 3.0 software (http://www.cs.umd.edu/hcil/hce/) divided the 67 unigenes into three major clusters (Fig 2.2; Table 2.6). Cluster I was divided into four subclusters Ia–Id, which contained unigenes highly expressed in the YL, YS, ML, and TB libraries, respectively. Clusters II and III showed high expression in roots: specifically, cluster II in the LR and TR libraries, and cluster III in the YR

library. Clusters Ia and Ic contained a number of photosynthesis-related genes, including chlorophyll-*a/b*-binding protein (Ia) and photosystem I reaction center subunit (Ic) (Table 2.6). Cluster II contained a unigene that encodes dihydroflavonol 4-reductase; this enzyme synthesizes leucoanthocyanidin, which is the direct precursor to (+)-catechin and (+)-gallocatechin. In cluster III, 10 out of 25 unigenes encoded stress-response proteins, including manganese superoxide dismutase and glutathione S-transferase.

2.3.4 Identification and analysis of EST-SSRs

An SSR motif search within the 5.3-k unigene set identified 1835 unigenes (34.9%) that harbored SSR motifs of ≥6 repeat units. Among these, the most frequent repeat motif was AG/TC, which was found in 24.4% of all unigenes, followed by AC/GT (6.5%) (Table 2.7).

I selected the 100 EST-SSRs with the highest numbers of repeat units and designed primer sets to amplify them (Table 2.8). Three of them (MSE0049, MSE0066, and MSE0089) had high homology to EST-SSRs reported by Sharma et al. (2009), but the other 97 were novel.

I tested the 100 EST-SSRs for their ability to amplify fragments within 16 tea accessions (Table 2.1). Of these, 71 produced well-amplified fragments, and 70 revealed polymorphism among the 16 accessions (Table 2.9). For 61 markers, only one or two fragments were amplified in each accession; these were considered single-locus markers. For the other 10 markers, some accessions had three or more amplified fragments; these were considered multi-locus markers. The single-locus markers had 1 to 15 alleles per locus, with an average of 8.2. Observed heterozygosities (H_O) ranged from 0 to 1.0, with an average of 0.64. Expected heterozygosities (H_E) ranged from 0 to 0.91, with an average of 0.72. Polymorphism information content ranged from 0 to 0.90, with an average of 0.69.

I investigated the transferability of the EST-SSRs to 14 *Camellia* species (Table 2.2). Of the 71 markers usable in *C. sinensis*, 70 were amplified in more than one species (Table 2.9). The average proportion of *C. sinensis* markers amplified in each of the 14 species was 87.1%. In *C. irrawadiensis*, a member of the same subgenus (*Thea*)

as *C. sinensis*, 68 markers (95.8%) were amplified (Table 2.10).

2.3.5 Identification and analysis of genomic SSRs

I designed a total of 733 primer sets from 1440 clones from SSR-enriched libraries. I used 30 of them to survey polymorphisms in the 16 accessions used in the analysis of EST-SSRs. PCR amplification was observed in 23 primer sets, which were used for fragment analysis using a DNA sequencer. Six primer sets showed many non-specific bands or stutter bands and were removed from further analysis. I analyzed the genotypes of the remaining 17 markers (Table 2.11).

2.4 Discussion

Before this study, the NCBI GenBank database held 14,246 ESTs and 34.5 $\times 10^6$ RNA-seqs from *C. sinensis*. This study identified 17,458 ESTs from seven cDNA libraries. Within the 5.3-k unigene set developed here, 732 unigenes had no significant matches by BLASTN homology searches against the tea ESTs and assembled sequences from RNA-seqs already deposited in GenBank, indicating that these unigenes are novel sequences. The lengths of 64.1% of the sequences in the 5.3-k unigene set were >500 bp, whereas in the unigenes generated by RNA-seq analysis, only 17.9% were longer than 500 bp. In general, EST analysis using Sanger sequencing generates longer sequence reads than RNA-seqs generated using a high-throughput Illumina GA IIx sequencer, so the difference in unigene length distribution can be attributed to the difference in sequencing technique.

The data presented here will provide a useful resource for research aimed at understanding physiological processes important for tea cultivation and quality, such as N assimilation and amino acid metabolism. In Japan, large amounts of N fertilizers are used in tea plantations, causing pollution of groundwater, rivers, and lakes. To decrease N fertilizers, it is important to develop tea cultivars with high N use efficiency (Tanaka and Taniguchi 2007). Therefore, I searched for unigenes related to N assimilation and found several that were homologous to genes for enzymes related to N assimilation, such as glutamine synthetase, glutamate dehydrogenase, ammonium transporters, and nitrate transporters (Table 2.5). In addition, the unigene set contains genes related to the

metabolism of 2-oxoglutarate, a key component of the interaction of nitrogen and carbon metabolism.

In addition to N compounds such as amino acids, secondary metabolites such as catechins and caffeine are important for tea quality. I found several unigenes related to the synthesis of these compounds. The metabolism of N compounds and secondary metabolites is regulated by environmental conditions. For example, in young tea leaves, high light intensity increases levels of catechins (Saijo 1980), but shading increases total N content and theanine content (Anan and Nakagawa 1974, Karasuyama and Matsumoto 1988). It will be important to decipher the mechanism of photoresponsive regulation of genes related to the metabolism of N compounds and secondary metabolites to enable improvement of these traits. Two unigenes related to photoresponse found among our ESTs will provide tools to analyze the associated regulatory mechanisms (Table 2.5).

Tea is well known as an Al-accumulator that grows well in very acidic soils containing high levels of Al^{3+}. This is of interest because Al toxicity limits the growth of many species in acidic soils (Morita et al. 2004, 2008), and the Al in the xylem sap of tea is complexed with citrate (Morita et al. 2004). I found three unigenes potentially related to Al response: one for citrate synthetase and two for Al-response proteins (Table 2.5). Further analyses, such as expression analysis of the response of tea to Al, might reveal whether these genes have roles in Al resistance or response.

Using the EST data, I performed digital northern analysis to identify unigenes with differential expression among different organs; 67 such unigenes were identified out of a sample of 144. Cluster analysis showed that the groups of unigenes highly expressed in each organ were related to different physiological functions. For example, several photosynthesis-related genes were highly expressed in the young leaves and mature leaves libraries. Cluster III, which showed high expression in the young roots library, was the largest cluster (28 unigenes). This result indicates that the physiological and developmental status of young roots is considerably different from that of other organs. Interestingly, dihydroflavonol 4-reductase (DFR) was highly expressed in tap roots and lateral roots (Fig. 2.2). Although catechins are not present in tea roots (Forrest and Bendall 1969), leucoanthocyanidin, which is the product of DFR and a precursor of (+)-catechin, is present. Thus, I assume this DFR in roots to be involved not in catechin

biosynthesis, but in other metabolic processes such as lignin or anthocyanin biosynthesis. One more unigene encoding DFR was found in the 5.3-k unigene set. This unigene was expressed in young stem tissue, and the sequence similarity between the two DFRs was 52%. I think that the DFR from young stem tissue is involved in catechin biosynthesis.

The results of this study could greatly increase both EST-SSRs and genomic SSRs. The hundreds of SSR markers developed here will enable genome-wide analysis such as the construction of linkage maps.

Ellis and Burke (2007) surveyed EST data from 33 species and showed that 2.5% to 21.1% (9.0% ± 0.1%, mean ± SEM) of unigenes contained SSRs. In contrast, the percentage of SSR-containing tea unigenes (34.9%) is high.

The proportion of multi-locus markers in the present study was higher than that reported by Sharma et al. (2009). I used a capillary sequencer for fragment analysis, whereas Sharma et al. (2009) used autoradiography of polyacrylamide gel electrophoresis gels, which has lower resolution. Thus, the difference in the proportion of multi-locus markers might have been caused by the difference in method. Because of the paleopolyploidy of *C. sinensis* (Shi et al. 2010), it is not surprising that the set of EST-SSRs reported here holds many multi-locus markers.

The 16 accessions used in this study include major tea cultivars in Japan, parental cultivars, and several foreign germplasms. These materials are representative of the genetic diversity of breeding materials in Japan. The EST-SSRs and genomic SSRs developed in this study were highly polymorphic among the 16 accessions. They should prove very useful for the construction of linkage maps, analysis of genetic diversity, and cultivar identification.

Most of the EST-SSR markers developed here are applicable to other *Camellia* species. Species other than *C. sinensis* contain useful traits that have been utilized in tea breeding; for instance, a parental line containing a high level of anthocyanins (Ogino et al. 2005) and a caffeine-less tea plant (Ogino et al. 2009) were developed from interspecific crosses. EST-SSR markers will enable genetic analysis of important agronomic traits of various *Camellia* species, thus expanding the usefulness of these species in tea breeding.

In conclusion, the tea ESTs obtained in this study are valuable resources for the analysis of gene function and for development of SSR markers. The 5.3-k tea unigene set contains novel transcripts from tea, and 67 out of 144 unigenes tested showed specific expression patterns among seven organs. The EST-SSR markers and genomic SSR markers developed in this study are highly polymorphic in *C. sinensis*, and the EST-SSRs are transferable to many other *Camellia* species. The SSR markers will support genetic studies in tea, such as the genetic mapping of agronomically important genes and analysis of genetic diversity.

Table 2.1. *Camellia sinensis* accessions used to investigate polymorphisms of EST-SSR loci.

Accession	Derivation	Origin	ID[a]
Sayamakaori	selected from seedlings of Yabukita	Japan	27029293
Kana-Ck17	introduced from Keemun, China	Japan	27001948
Minamisayaka	Miya A6 × NN27	Japan	–
Yabukita	selected from indigenous seedlings in Japan	Japan	27027257
Shizu Inzatsu 131	selected from hybrids of var. *sinensis* and var. *assamica*	Japan	–
Asatsuyu	selected from indigenous seedlings in Japan	Japan	27027248
Miyamakaori	Kyo Ken 283 × Saitama No. 1	Japan	–
ME52	selected from indigenous seedlings in Japan	Japan	27025724
Shizu Zai 16	selected from indigenous seedlings in Japan	Japan	–
Shizu 7132	selected from seedlings of Yabukita	Japan	–
KaCp1	introduced from Pingshui, China	China	–
Z1	selected from seedlings of Tamamidori	Japan	–
Benifuki	Benihomare × Makura Cd86	Japan	–
Ak1699	introduced from Darjeeling, India	India	27002929
Makura No. 1	introduced from India	India	27003028
Taiwan Yamacha 95	introduced from Taiwan	Taiwan	27003335

[a] Accession ID of the NIAS Genebank.

Table 2.2. *Camellia* species used to investigate transferability of EST-SSRs.

Name of accession	Species	Subgenus
Taliensis Midorime	*C. taliensis*	*Thea*
Irrawadiensis	*C. irrawadiensis*	*Thea*
Suzukayama	*C. japonica*	*Camellia*
Pitardii	*C. pitardii*	*Camellia*
Hongkongensis	*C. hongkongensis*	*Camellia*
Chekiangoleosa	*C. chekiangoleosa*	*Camellia*
Saluenensis	*C. saluenensis*	*Camellia*
Kissi	*C. kissi*	*Camellia*
Oleifera	*C. oleifera*	*Camellia*
Sasanqua Matsumoto 1	*C. sasanqua*	*Camellia*
Furfuracea	*C. furfuracea*	*Camellia*
Cuspidata	*C. cuspidata*	*Metacamellia*
Salicifolia	*C. salicifolia*	*Metacamellia*
Granthamiana	*C. granthamiana*	*Protocamellia*

Table 2.3. cDNA library statistics.

Source of RNA	No. of clones	ESTs	Unique transcripts
young roots	3,072	4,529	1,587
tap roots	2,880	1,927	1,013
lateral roots	2,880	2,316	1,230
young leaves	2,880	2,233	1,090
terminal buds	2,880	2,221	1,066
young stems	2,880	2,147	1,187
mature leaves	2,880	2,085	846
Total	20,352	17,458	5,262[a]

[a] Number of unigenes generated from 17,458 ESTs.

Table 2.4. Tea plant EST summary.

Feature	Value
Sequence information	
Total number of sequences	17,458
Total nucleotides (bp)	8,391,523
Average read length (bp)	481
GC content (%)	44.0
Unigene information	
Number of scaffolds	442
Number of contigs	1,851
Number of singletons	2,969
Number of sequences in unigenes	
2 ESTs	958
3–5 ESTs	823
6–10 ESTs	335
11–15 ESTs	83
≥16 ESTs	94
Unigene length distribution	
Length	No. of unigenes
100–500 bp	1,890
501–1000 bp	2,900
>1000 bp	472

Table 2.5. Unigenes related to important biological processes in tea.

Classification	Function	No. of unigenes	No. of ESTs
Aluminum response	aluminum-induced protein	2	17
	citrate synthetase	1	1
Caffeine biosynthesis	caffeine synthase	1	9
Catechin biosynthesis	4-coumarate CoA: ligase	2	5
	anthocyanidin reductase	2	6
	chalcone isomerase	4	16
	chalcone synthase	3	20
	cinnamate 4-hydroxylase	2	3
	dihydroflavonol 4-reductase	2	41
	flavonoid 3'-hydroxylase	4	5
	flavonol synthase	3	9
	leucoanthocyanidin reductase	1	2
	phenylalanine ammonia-lyase	2	4
Nitrogen assimilation and amino acid metabolism	2-oxoglutarate malate translocator	2	2
	alanine aminotransferase	3	3
	amino acid channel protein	1	1
	amino acid transporter	4	4
	ammonium transporter	2	2
	aspartate aminotransferase	2	2
	glutamate dehydrogenase	1	2
	glutamate synthetase	1	1
	glutamine dumper	1	1
	glutamine synthetase	3	12
	glycine decarboxylase	3	7
	NAD^+-dependent isocitrate dehydrogenase	1	1
	$NADP^+$-dependent isocitrate dehydrogenase	1	2
	nitrate transporter	1	1
	serine hydroxymethyltransferase	1	1
Photoresponse	cryptochrome	1	1
	CIP8 (COP1-interacting protein 8)	1	1

Table 2.6 The unigenes which showed differential expression in seven organs.

Cluster	Unigene	Annotation	GO terms
Ia	ug2071	chlorophyll a/b binding	membrane;plastid;generation of precursor metabolites and energy;photosynthesis
Ia	ug2085	chlorophyll a/b binding protein	generation of precursor metabolites and energy;photosynthesis;membrane;thylakoid;binding;protein modification process;plastid
Ia	ug2001	light-harvesting complex II protein Lhcb3	membrane;plastid;generation of precursor metabolites and energy;photosynthesis
Ia	ug2083	serine hydroxymethyltransferase	binding;cellular amino acid and derivative metabolic process;transferase activity;metabolic process;cellular process
Ia	ug2087	alpha tubulin 1	cellular process;C;nucleotide binding;cellular component organization;structural molecule activity;hydrolase activity;cytoskeleton
Ia	ug1989	no annotation	
Ia	ug2024	photosystem II protein Z	membrane;thylakoid;generation of precursor metabolites and energy;photosynthesis;plastid
Ia	ug2050	no annotation	
Ia	ug2058	chlorophyll a/b binding	membrane;plastid;generation of precursor metabolites and energy;photosynthesis
Ia	ug2054	photosystem II 10 kDa	membrane;thylakoid;plastid;photosynthesis
Ib	ug2033	no annotation	
Ib	ug2067	metallothionein-like protein	binding
Ib	ug2066	non-specific lipid transfer protein precursor	lipid binding;transport
Ib	ug2080	non-specific lipid transfer protein precursor	lipid binding;transport
Ib	ug2006	unknown protein	cytoplasm;plastid
Ib	ug2007	no annotation	
Ic	ug2061	photosystem II CP43 protein	thylakoid;plastid;membrane;generation of precursor metabolites and energy;photosynthesis;intracellular;F;protein modification process;binding
Ic	ug2091	hypothetical ctc00065-like protein	mitochondrion;plastid
Ic	ug2064	photosystem I reaction center subunit	membrane;thylakoid;plastid;photosynthesis
Ic	ug2089	no annotation	
Ic	ug2090	histone H3	intracellular;DNA binding;cellular component organization;cellular process;plastid;nucleus
Id	ug2018	unknown protein [Camellia sinensis]	
Id	ug2021	extensin like protein	transport
Id	ug2002	unknown protein [Camellia sinensis]	
Id	ug2079	histone H4	intracellular;DNA binding;cellular component organization;cellular process;protein binding;nucleus;cytoskeleton
Id	ug2029	PREDICTED: hypothetical protein	
Id	ug2082	histone H4	intracellular;DNA binding;cellular component organization;cellular process;protein binding;nucleus;cytoskeleton
Id	ug1961	histone H3	intracellular;DNA binding;cellular component organization;cellular process;plastid;nucleus
II	ug2055	lipid binding	transport
II	ug2060	unknown protein	
II	ug2015	no annotation	
II	ug2084	no annotation	
II	ug2078	dihydroflavonol 4-reductase	metabolic process;cellular process;binding;catalytic activity
II	ug2092	cell wall-associated hydrolase	kinase activity;hydrolase activity
II	ug2074	H^+-transporting two-sector	vacuole;transport;generation of precursor metabolites and energy;nucleobase, nucleoside, nucleotide and nucleic acid metabolic process;biosynthetic process;membrane;hydrolase activity;transporter activity;plasma membrane
II	ug2093	unknown protein	binding
II	ug2088	unknown protein	metabolic process;catalytic activity
II	ug2065	specific tissue protein 2	
II	ug2075	organ-specific protein	

Table 2.6 Continued

Cluster	Unigene	Annotation	GO terms
III	ug1784	unknown protein	binding;catalytic activity
III	ug1016	40S ribosomal protein S5	structural molecule activity;ribosome;translation
III	ug1854	isopentenyl diphosphate isomerase	hydrolase activity;biosynthetic process;cellular process;lipid metabolic process;catalytic activity
III	ug2039	manganese superoxide dismutase	binding;metabolic process;response to stress;cellular process;mitochondrion;catalytic activity
III	ug2076	GEG protein	extracellular region;response to stress;cell wall
III	ug1560	glutathione S-transferase	response to endogenous stimulus;signal transduction;transferase activity
III	ug2053	germin-like protein 6	response to endogenous stimulus;signal transduction;extracellular region;cell wall;receptor activity;F;binding
III	ug1822	nucleoside diphosphate kinase	kinase activity;response to external stimulus;signal transduction;response to abiotic stimulus;biosynthetic process;nucleobase, nucleoside, nucleotide and nucleic acid metabolic process;nucleotide binding;binding
III	ug2038	PREDICTED: hypothetical protein	
III	ug2072	early response to dehydration 15-like protein	
III	ug0640	lipid binding	structural molecule activity;transport
III	ug2046	calcium-binding EF hand family protein	binding
III	ug1974	alcohol dehydrogenase	catalytic activity;cytoplasm;metabolic process;binding
III	ug2063	fructose-bisphosphate aldolase	carbohydrate metabolic process;generation of precursor metabolites and energy;catabolic process;catalytic activity
III	ug2020	pathogenesis-related protein 10	
III	ug1701	hypothetical protein	
III	ug2045	60S ribosomal protein	ribosome;structural molecule activity;translation
III	ug2081	no annotation	
III	ug1557	60S ribosomal protein l35a	structural molecule activity;cellular process;ribosome;cytosol;translation
III	ug1965	60S acidic ribosomal protein P0	ribosome;structural molecule activity;cellular process;translation
III	ug2056	histone H2b	intracellular;DNA binding;cellular component organization;cellular process;nucleus
III	ug2086	no annotation	
III	ug2048	peroxidase 72	binding;metabolic process;catalytic activity;F;response to stress
III	ug2049	germin-like protein 6	response to endogenous stimulus;signal transduction;extracellular region;cell wall;receptor activity;F;binding
III	ug2047	no annotation	

Table 2.7. Number and motif distribution of EST-SSRs.

Motif	Number of unigene sequences containing the number of repeats specified			
	≥6 repeats		≥10 repeats	
	n	(%)	*n*	(%)
AG/TC	1,284	24.4	608	11.0
AT/TA	271	5.2	127	2.3
AC/GT	344	6.5	156	2.8
GC/CG	17	0.3	13	0.2
AAC/GTT	45	0.9	24	0.4
AAG/CTT	88	1.7	42	0.8
ACC/GGT	120	2.3	60	1.1
ACG/CGT	26	0.5	15	0.3
ACT/AGT	57	1.1	24	0.4
AGC/GCT	37	0.7	24	0.4
AGG/CCT	81	1.5	35	0.6
ATC/GAT	52	1.0	28	0.5
TAT/ATA	24	0.5	13	0.2
CGC/GCG	23	0.4	10	0.2
Any SSR motifs[a]	1,835	34.9	878	16.0

[a] Number of unigenes containing any SSR motifs listed in this table.

Table 2.8. List of EST-SSR markers used in this study

marker name	repeat region	fw primer for postlabel	rv primer for postlabel	accession no.
MSE0019	(ac)23(tc)12, (ac)3, (ac)4	accatcaaactactctcatcccca	gtttcccacttctatgtcctttgccat	FS951295
MSE0020	(tc)21,(tc)3(ta)7,(ac)3	attcccacttgacaccaaaccaag	gtttctccggtcacctctatcttgtcc	FS949264
MSE0021	(tc)20(ta)10,(ta)3	accatccatccataccattcttca	gtttatggaatgcatccaacaacaagg	FS951578
MSE0022	(tc)19(ta)9, (ac)5	actgcagtgtagaaaagccccaat	gtttcgatctgggagcttcttggagata	AB361048
MSE0023	(tc)13, (tc)7, (ct)3, (ta)5	agaaacaaccacctcccttgatct	gtttagcagcgaagaaggattcattacg	AB485966
MSE0024	(ag)3, (ag)11, (gaa)3, (ag)4, (ag)3, (ag)3	ataaccatcattccaaagccacaa	gtttccgatttcgacatctcttcgtctt	FS951122
MSE0025	(tc)5, (tc)5, (tc)14, (ag)3	atatatcctccctcaggctcctcc	gtttcagcaatgaaacgaaacccagatt	FS948413
MSE0026	(ac)3, (ag)7, (ac)4, (ag)3, (ag)6, (tc)3	atcaccaaaggaaacccctactgg	gtttcgcaggtggatctggagaaaatac	AB461371
MSE0027	(cct)3,(tc)19,(tc)3	atccctcttacttcaatcgacccc	gtttgaagggcggtacttgtatggat	FS952009
MSE0028	(tc)21,(ag)3	acaacccaaaatttcacatgatagc	gtttcaacaacatcgctagaggtcca	FS951387
MSE0029	(ag)14, (ag)7, (gga)3	atagccaatcaagctcctcctcct	gtttctgttcctcccttgatgatcg	FS949897
MSE0030	(tc)4, (ca)3, (tc)5, (tc)11	atccaaaaccctagtttcactcca	gtttacgctctgtatcggtgaaggctac	AB461369
MSE0031	(tc)8,(tct)9,(ac)3,(ca)3	atgatctgaaggtgaagggtttcg	gtttcttatgagacttggccttcgc	FS951496
MSE0032	(ag)4,(ac)13,(tc)3,(tc)3	attttccctcaaaccttcttcagc	gtttaatcccctggtgctcattctgt	FS949441
MSE0033	(tc)4,(tc)12,(tg)3,(tc)3	acaaactcaaaccatctcagcggt	gtttcagagataatggctcgttggga	FS951827
MSE0034	(tc)14,(tc)4,(tc)4	agctctctgttccgaagcttttag	gtttcacctccaaacaaccaaaac	FS951853
MSE0035	(tc)13, (ag)4, (ag)5	atccttcaatcttccattgacgc	gtttcacaaaaaccccaaaataccctcc	AB461372
MSE0036	(ag)11,(ag)8,(tc)3	atgtgctattgttggttgtggacc	gtttaaaagcttggtgaaagatagcccc	FS951471
MSE0037	(ca)3, (tc)12, (ag)4, (gaa)3	attacccatttgccttgcttggtc	gtttgaagcaaatttgggaaatcagtgc	AB485967
MSE0038	(ag)5, (at)3, (tg)3, (ta)3, (ag)8	attcctatgtctgagcccttcatt	gtttgtaggcaagctgaacatcgtctc	FS950252
MSE0039	(ag)21	acattctgctgcaattacacatcc	gtttcttttggtaagcggctcattgtc	AB485968
MSE0040	(tc)18, (ctt)3	actccgacgagtccagttcctaac	gtttgtacatttgcagaggcagagc	AB485969
MSE0041	(tc)4,(tc)4,(tc)10,(tc)3	agacagtgctccatctccatctcc	gtttacccttggattttgcccacatac	FS950273
MSE0042	(tc)16, (tc)5	agactcgctcgctctcaacaaac	gtttaagcggagataatggtggttagca	FS950639
MSE0043	(ag)4, (ta)7, (ag)10	atagggcagagatcagctggaaaa	gtttccatcttgctgagttgttctgc	FS951029
MSE0044	(ta)11, (ag)4, (ag)6	atgcttagcagccatcattcaaca	gtttctttttacgaacaacacccgc	FS948949
MSE0045	(ag)14, (ga)3, (ag)4	atggcgataacttcgagacacaaa	gtttccctccctccctttttaatggtt	AB461368
MSE0046	(tc)14,(tc)3,(ct)3	actccctcacatcactctcagtcg	gtttgagacagtgtgagtctatcgggg	FS950610
MSE0047	(tc)13, (tc)4, (tc)3	agatcatctctaaaccccctaaaaccc	gtttgatgttggaatggttgaaggaga	AB485970
MSE0048	(tc)9,(tct)3,(tc)4,(tc)4	ataaagccctaaaccctcgcaact	gtttgcctggaaccttgggaactttaac	FS948085
MSE0049	(at)3, (ag)14, (ag)3	atgtagcattttggcttattggatcac	gtttcctgtcaaagctcctcttgtgg	AB361047
MSE0050	(ttc)3, (tc)14, (ca)3	attgagtccgtcattgttccacac	gtttcttttgatcgagggagtcagagga	FS947935
MSE0051	(tc)15, (ta)4	acttgcaagtggacaaaactgttc	gtttagctgcagagacgtaggcatttct	AB461365
MSE0052	(ac)3, (tc)4, (tc)9, (tc)3	agttacggtacttcttgttagctggc	gtttggttggtgtgtctggtgattctt	AB505872
MSE0053	(ga)3, (tc)16	ataatccccaaaatttgtctcagtg	gtttcacgatgagagtgtgaccgtt	AB461367
MSE0054	(ta)4, (tc)15	atacaagaacgaaacaactctccg	gtttgagagctggaatgggtgtctgagt	FS947912
MSE0055	(cca)3, (acc)3, (cca)3, (ta)6, (ta)4	atcacgaaccacctcctcctcat	gtttgcagatatacataactccaaccct	FS949373
MSE0056	(tc)10, (ctc)3, (aga)3, (ag)3	atcccatcccatcccatatcatc	gtttcaacaacatcgctagaggtcca	AB361052
MSE0057	(tc)4,(tc)12,(at)3	atctctctctcacgtttccgttta	gtttggactcatcagcagcaaggata	FS951743
MSE0058	(tc)12, (tc)3, (tc)4	atctttatcttcatcgccgaccag	gtttagatttggatgctcaatgaacggt	FS951427
MSE0059	(ag)3, (ag)5, (gaa)11	attccgtgaggagacgacgaag	gtttgactttgtggtggatctttcc	FS950702
MSE0060	(tc)5,(tc)5,(tc)5,(tc)4	atttatcgaaacgtgtctccgtca	gtttaacacccgaatctccaataagcaa	FS948154
MSE0061	(tc)3, (tc)3, (tc)9, (ac)3	acacacatacaacccaacccaagc	gtttggaggactctgcatccattacagc	FS949588
MSE0062	(ag)18	acagagaacattgaacatcaagcc	gtttaggttctttccccaatcgacct	AB461364
MSE0063	(tgt)3, (ag)9, (aga)3, (ag)3	acattctagagggagaaaaagtgtga	gtttcccagctctgtaaaacgaccac	AB505868
MSE0064	(ac)3,(ag)7,(ag)4,(ag)4	accagttaccaagccctcaaatca	gtttccacagtgagttccacatccatt	FS949747
MSE0065	(tc)3,(tc)15	agtcaccctaaccaaaccaaatca	gtttcctgtaacatcatggtggctg	FS949036
MSE0066	(tc)4, (tc)4, (tc)3, (tc)4, (tc)3	ataaagccagtggccaacactgc	gtttgcagcaagatcagtcatgt	FS949044
MSE0067	(tc)18	atcttaagtctcccctcttgcgct	gtttggggattgagaggaggaatgtaa	FS947993
MSE0068	(tc)9, (tc)5, (tc)3	acgtttattctcttctcgctcagt	gtttgaggactcgaagaaattgaagcca	AB505871
MSE0069	(tc)17	atactcaccgcactctctcccact	gtttcagtggcaggtggtgaaactgtag	FS952134
MSE0070	(tcc)3(tc)8,(tc)6	atcacctctcatcacccttttctca	gtttgcattcatggttgcaagagagaga	AB505870
MSE0071	(tc)3, (tc)14	atcctccaccgaagtccgatcc	gtttagaagttgttcctcaggccgaa	AB505867
MSE0072	(tc)14, (tc)3	attcagtccaccctacccccctaaa	gtttagcatcctttggataccctttctgg	AB361049

Table 2.8. continued

marker name	repeat region	fw primer for postlabel	rv primer for postlabel	accession no.
MSE0073	(tc)14t(ac)3	attcggctctctctccctctct	gtttaacgaaggaacaagaagaagcacg	FS947738
MSE0074	(tc)7, (ct)3, (tc)3, (ct)3	acaatattccactcccccaaacac	gtttgccctcttgttcttcttcttgtcc	FS948120
MSE0075	(tc)4, (tc)12	acacgttttgtagcctaggttcttca	gtttaaatcaatgtttcgatcgctgctc	FS949483
MSE0076	(ca)3, (ac)4, (tat)3, (gta)3, (tgg)3	acatccattcatcaacaagggaat	gtttgcttgatttcccagctgttctt	AB505869
MSE0077	(tc)3, (tc)10, (cg)3	accaggcgtctttgtatacgtctc	gtttcgaaatgcactctctcattttcg	AB361050
MSE0078	(ag)16	actgcgtccttctttccttgcac	gtttggagtcaattagggctgggttt	AB461366
MSE0079	(ag)16	agtatttcacttttgcagagccca	gtttagctttgaagataggttggcgatg	AB505866
MSE0080	(ac)4, (ta)4, (tca)3, (tc)5	ataccctcctcttcccatcacctt	gtttaccagaacagaggacaggagcag	FS949599
MSE0081	(tc)4, (tct)3, (tc)9	atactttccccctccttcctcctt	gtttgttatccaccggtaaagcca	FS950583
MSE0082	(ca)3, (ta)13	atcaaaatccctcccccaaaaa	gtttcaaactaagaaccaatgccacca	FS948901
MSE0083	(tct)4(ct)3,(tc)3,(tct)6	atctgatcgaaatcaccaatgcct	gtttgaagaagagagagaagtccgtgcg	FS951085
MSE0084	(tc)16	atgcttttcttcccacctctttcg	gtttgacgctccattcgaataagttgct	FS949655
MSE0085	(tc)3,(tc)4,(tc)9	atgtcacgaacatggcaccttt	gtttcgcctgacacttcttcgtgatt	AB505865
MSE0086	(tc)16	attcctctatcttccctccccaaa	gtttagggtatgatcgcaacttcttcca	FS951412
MSE0087	(ag)3, (ag)13	attttattcctcctactgcgccga	gtttaaagccttgacgcgagacttctta	AB505873
MSE0088	(ga)3, (tg)3, (tc)9	accaaacaaaggctagaccagaatg	gtttgtctttgaagggtttcgttg	FS951504
MSE0089	(ag)5, (ag)7, (tg)3	accgtcttcttcttcttctctccc	gtttaatcacctggaaaacagcggaa	FS949251
MSE0090	(tc)15	acgaaccacaggaaacaatctcaa	gtttcttttctgcttcctccacagcatc	FS950145
MSE0091	(ta)7,(ag)8	acgtcgttctttctgagattttgt	gtttaagtcaacatgaagccaacccaat	FS949737
MSE0092	(ca)3,(ag)12	actcatacaagccttcaagcacgc	gtttcgtccctcctcctgaagaact	FS949294
MSE0093	(ac)4,(tc)11	actgcatacacaccttgcactgaa	gtttgaggacactgacatggactgga	FS948326
MSE0094	(ta)4, (ta)8, (ta)3	agaaaaaccattctgcacaccact	gtttatggaaggaacttggaggttgga	FS948521
MSE0095	(cct)9,(cgc)3,(gct)3	agattcatcgattgctttatctccc	gtttagcctatcacccttcattgttgga	FS948314
MSE0096	(tc)12, (ta)3	agctctttccagcttcacctcata	gtttacacaaatatgcaccatcaccgtt	FS951767
MSE0097	(ga)3,(cca)3,(ag)9	aggactgaaggagtagaaaatcca	gtttcatgatgtaggctccaccgatt	FS951041
MSE0098	(cca)5, (cca)4, (cca)6	atcccatctccatctccaaattca	gtttgcaacaaagcttcaaactcct	FS951011
MSE0099	(ac)3, (ag)3, (ag)3, (ag)6	atctcagcttccttctttacaaacaca	gtttatcaccggtggcaaaggtatg	FS952164
MSE0100	(tc)15	atctttccgtgtacatacaccccc	gtttgaattgttggaggccgtagaattg	AB461370
MSE0101	(gc)3, (aac)3, (cca)6, (ccg)3	atgctctcctctgtttacgccttg	gtttacgaattctgaccacgaaacccta	FS950763
MSE0102	(tc)5(ctc)3, (tc)4, (tc)3	atttcctccaccatgcggttt	gtttagaacggcagtgcttaggtttttg	FS951686
MSE0103	(tc)14	acaagcaaaggcattacaaaaagc	gtttcaccaactgaaagaatccccat	FS951354
MSE0104	(ag)14	acacaaccactctgcaccagacac	gtttccggattcagtcagaaatttggag	FS949898
MSE0105	(cgc)3, (cca)6, (cca)5	acatccaaaaacaaggaagctcca	gtttgttgttggagattgaggtggtg	FS949892
MSE0106	(tc)4, (tc)6, (ta)4	acgcttctcttcctttcatctcaa	gtttcaacaaagaagacagccaaagcct	FS947853
MSE0107	(tc)8, (ca)3, (cct)3	actctctactcctgcgcaatctca	gtttcaaagatgttgctctcgtcaacc	AB485971
MSE0108	(tc)6,(ta)8	agtccatggtggttgatgatcctt	gtttgggagtaggattcttgcagagc	FS948805
MSE0109	(tc)10, (tg)4	ataaaagacaataccatcccaggaa	gtttccagaagtgagaagggttgtcg	FS948318
MSE0110	(tc)11, (cac)3	atataatcgcggtttccactcact	gtttggattcatggctattgaaggcttg	FS950817
MSE0111	(tc)11, (ct)3	atatcagaagtgtggagagtgcct	gtttaccatttgaggaatgcgtgacttt	FS951008
MSE0112	(ag)5, (ag)3, (ag)3, (tg)3	atcagagtttgccacatggacagt	gtttcatcttctgcttgtcccccttc	FS947995
MSE0113	(tc)14	atccttctgcaactccagcaatcc	gtttgagattgaccatctttcatcgga	AB485972
MSE0114	(tc)14	atgcctgtactcttccattcttgc	gtttcttttttcagtcactagctctggggc	FS948705
MSE0115	(tc)3, (tc)4, (tc)7	attcacctagccctcgcatttcta	gtttgtcatctccatcatccattcca	FS948956
MSE0116	(tc)10, (tc)3	acacacttcacgctgcactttact	gtttgtcgcagtaaaccttaccctc	FS948104
MSE0117	(aca)3, (ag)4, (ag)3(tg)3	acacatggtaacctcattacttagctc	gtttgaccaatggacggacaatatgac	FS948642
MSE0118	(tc)4, (ag)9	acacctattaagcccaccgctctt	gtttaactgtagctgcacctggaacctc	FS949980

Table 2.9. Features of EST-SSRs and polymorphism information in 16 tea accessions.

Marker name	SSR motif	Position of repeat motifs [a]	Approx. size range (bp)	No. of accessions w/ amplification	No. of transferable species	No. of loci[b]	No. of alleles	Heterozygosity[c]		PIC value[d]
								H_E	H_O	
MSE0019	(ac)23(tc)12	5′	105–150	16	13	m				
MSE0021	(tc)20(ta)10	5′	265–310	13	7	s	10	0.80	0.46	0.78
MSE0022	(tc)19(ta)9	5′	165–210	16	14	s	14	0.89	0.88	0.88
MSE0023	(tc)13, (tc)7	unknown	180–240	16	12	s	8	0.83	0.63	0.81
MSE0024	(ag)11	5′	255–285	16	14	s	9	0.82	0.56	0.80
MSE0025	(tc)14	unknown	260–305	16	14	m				
MSE0026	(ag)7, (ag)6	5′	275–300	16	14	s	7	0.79	0.56	0.75
MSE0027	(tc)19	5′	105–135	15	6	s	5	0.56	0.47	0.52
MSE0029	(ag)14, (ag)7	5′	365–340	16	11	m				
MSE0030	(tc)11	5′	245–270	16	14	s	10	0.78	0.81	0.75
MSE0035	(tc)13	5′	210–250	16	13	s	10	0.82	0.81	0.80
MSE0037	(tc)12	tr, 3′	200–250	16	14	m				
MSE0038	(ag)8	5′	300–320	16	13	m				
MSE0039	(ag)21	unknown	135–180	16	14	m				
MSE0040	(tc)18	5′, tr	125–155	16	14	s	9	0.73	0.75	0.71
MSE0042	(tc)16	unknown	100–115	16	12	s	8	0.77	0.50	0.74
MSE0043	(ta)7, (ag)10	unknown	170–210	16	12	s	10	0.79	0.69	0.77
MSE0044	(ta)11, (ag)6	5′	120–145	16	10	s	9	0.75	0.56	0.72
MSE0045	(ag)14	unknown	215–230	16	14	s	6	0.75	0.88	0.71
MSE0047	(tc)13	5′	245–275	16	14	s	11	0.79	0.75	0.77
MSE0049	(ag)14	unknown	220–250	16	14	s	10	0.84	0.81	0.83

Table 2.9. continued

Marker name	SSR motif	Position of repeat motifs [a]	Approx. size range (bp)	No. of accessions w/ amplification	No. of transferable species	No. of loci [b]	No. of alleles	Heterozygosity [c]		PIC value [d]
								H_E	H_O	
MSE0050	(tc)14	unknown	265–285	16	14	s	11	0.85	0.88	0.83
MSE0051	(tc)15	5′	185–215	16	14	s	11	0.87	0.81	0.86
MSE0052	(tc)9	tr	260–285	16	13	s	12	0.86	0.81	0.85
MSE0053	(tc)16	5′	250–275	16	12	s	11	0.83	0.81	0.81
MSE0054	(tc)15	5′	165–195	16	14	m				
MSE0055	(ta)6	unknown	235–265	16	14	s	3	0.22	0.25	0.21
MSE0056	(tc)10	5′, tr	220–240	13	13	s	6	0.75	0.77	0.71
MSE0058	(tc)12	unknown	185–275	16	14	m				
MSE0059	(gaa)11	5′	195–225	16	14	s	7	0.71	0.63	0.67
MSE0061	(tc)9	5′	135–165	16	13	m				
MSE0062	(ag)18	tr	110–140	16	14	s	9	0.84	0.81	0.82
MSE0063	(ag)9	5′, tr	230–255	16	14	s	10	0.79	0.81	0.77
MSE0066 [e]	(tc)4, (tc)4, (tc)3, (tc)4, (tc)3	5′, tr	240–260	16	5	s	1	0.00	0.00	0.00
MSE0067	(tc)18	5′	135–165	16	4	s	11	0.88	0.69	0.87
MSE0068	(tc)9	5′	255–275	16	14	s	10	0.80	0.88	0.77
MSE0069	(tc)17	unknown	195–225	15	11	s	11	0.85	0.40	0.83
MSE0070	(tc)8,(tc)6	5′	125–140	16	11	s	6	0.72	0.56	0.68
MSE0071	(tc)14	5′	110–135	16	13	s	8	0.80	0.63	0.78
MSE0072	(tc)14	5′, tr	295–320	16	14	s	9	0.81	0.75	0.79
MSE0074	(tc)7	unknown	200–210	16	13	s	4	0.62	0.56	0.54
MSE0075	(tc)12	unknown	140–200	16	14	m				

Table 2.9. continued

Marker name	SSR motif	Position of repeat motifs [a]	Approx. size range (bp)	No. of accessions w/ amplification	No. of transferable species	No. of loci[b]	No. of alleles	Heterozygosity[c]		PIC value[d]
								H_E	H_O	
MSE0076 [e]	(ca)3, (ac)4, (tat)3, (gta)3, (tgg)3	tr	240–245	16	13	s	3	0.36	0.19	0.33
MSE0077	(tc)10	5′	120–140	16	14	s	8	0.68	0.56	0.64
MSE0078	(ag)16	5′	140–160	16	14	s	7	0.80	0.69	0.77
MSE0079	(ag)16	5′	100–115	16	14	s	11	0.76	0.81	0.74
MSE0080 [e]	(ac)4, (ta)4, (tca)3, (tc)5	unknown	225–245	16	13	s	5	0.52	0.38	0.48
MSE0081	(tc)9	5′	100–140	16	8	s	9	0.82	0.31	0.79
MSE0082	(ta)13	unknown	155–185	16	13	s	8	0.81	0.81	0.78
MSE0083	(tct)6	5′	235–265	16	14	s	9	0.84	0.69	0.83
MSE0084	(tc)16	5′	395–420	16	14	s	11	0.83	0.81	0.81
MSE0087	(ag)13	5′	265–285	16	14	s	7	0.79	0.75	0.76
MSE0088	(tc)9	unknown	185–195	15	3	s	4	0.46	0.33	0.42
MSE0089	(ag)7	5′, tr	290–310	16	14	s	7	0.61	0.63	0.59
MSE0094	(ta)8	3′	180–220	16	14	s	8	0.70	0.56	0.67
MSE0096	(tc)12	3′	235–260	16	14	s	11	0.74	0.75	0.72
MSE0098	(cca)6	tr	245–270	16	14	s	7	0.62	0.81	0.58
MSE0099	(ag)6	5′	285–300	16	14	s	7	0.73	1.00	0.70
MSE0100	(tc)15	5′	235–260	15	7	s	10	0.86	0.93	0.84
MSE0101	(cca)6	5′	240–270	16	14	s	7	0.77	0.75	0.73
MSE0102 [e]	(tc)5(ctc)3, (tc)4, (tc)3	tr	305–345	16	12	s	9	0.78	0.63	0.76
MSE0103	(tc)14	5′	170–195	16	13	s	8	0.74	0.81	0.72
MSE0106	(tc)6	unknown	145–175	16	14	s	8	0.80	0.44	0.77

Table 2.9. continued

Marker name	SSR motif	Position of repeat motifs [a]	Approx. size range (bp)	No. of accessions w/ amplification	No. of transferable species	No. of loci[b]	No. of alleles	Heterozygosity[c]		PIC value[d]
								H_E	H_O	
MSE0107	(tc)8	5′, tr	290–315	15	14	s	10	0.78	0.87	0.76
MSE0108	(tc)6(ta)8	unknown	245–270	16	14	s	9	0.79	0.81	0.77
MSE0109	(tc)10	unknown	105–125	13	5	s	6	0.67	0.31	0.63
MSE0112 [e]	(ag)5, (ag)3, (ag)3, (tg)3	unknown	280–285	16	14	s	2	0.48	0.56	0.37
MSE0113	(tc)14	5′	350–380	16	0	s	9	0.84	0.75	0.82
MSE0114	(tc)14	unknown	195–205	16	2	s	5	0.61	0.94	0.53
MSE0116	(tc)10	5′	175–195	16	14	s	6	0.65	0.38	0.61
MSE0117 [e]	(aca)3, (ag)4, (ag)3(tg)3	tr, 3′	105–125	16	14	s	7	0.59	0.31	0.55

[a] 5′, 5′-UTR; 3′, 3′-UTR; tr, translated region.

[b] s, single locus; m, multi-locus.

[c] H_E, expected heterozygosity; H_O, observed heterozygosity.

[d] PIC, polymorphism information content.

[e] All SSR motifs in these markers are <6× repeats, but these markers were included in the analysis because the total number of repeats is >

Table 2.10. Amplified fragment size data of EST-SSRs in 14 Camellia species

Marker name	*Number of transferable species*	subgenus *Thea*		subgenus *Camellia*				
		C. irrawadiensis	*C. taliensis*	*C. japonica*	*C. pitardi*	*C. hongkongensis*	*C. chekiangoleosa*	*C. saluenensis*
MSE0019	13	109/136/140	112/125/129	131/151/169	114/145/156	140/145	112/149/158/163	-
MSE0021	7	-	269	278/282	278	-	-	265/278
MSE0022	14	175/189	189	180/186	186/192/198	176/194	186	184/211
MSE0023	12	197	197/201	196/204	184/197/200/212/227	179	-	200
MSE0024	14	257	259	265/276	245/259/263/270/274	268/282	257/259	256/266
MSE0025	14	291/295	280/287/291	272/280/288/295	270/295	285	285/293	293
MSE0026	14	294	289	272/285	287/293	297/317	285/287	297
MSE0027	6	107/109	109	-	121	110	-	-
MSE0029	11	382/384	377	-	364	384/390	373/380	377
MSE0030	14	250/255	255	259	239/248/257/261	248/255	247/252/263	252/266
MSE0035	13	233/237	235	246/254	218/220/231	215	222/233	218/237
MSE0037	14	189/214	174/184/227/239	172/185/221/232	214/228/247	186/197	214/231	164/231/237
MSE0038	13	297	297	304/322	301/314	310	-	303
MSE0039	14	140/144	144	150	148	144	144	148
MSE0040	14	151/160	147/151	138/149	142/153/157/161	138/165	160/161	142/147
MSE0042	12	101/105	103	116	101/105/108/110/112	112/114	103/107/114	105
MSE0043	12	189	174/184	172/185	-	186/197	192/203	164/176
MSE0044	10	134	132	111	-	138	127/130	-
MSE0045	14	229	229	226/241	208/210/225	229/235	223	234/235
MSE0047	14	275	278/282	262/264	259/273	264	275/288	268/273
MSE0049	14	225/230	228/232	220/222	202	210/238	211/220	202
MSE0050	14	273/276	268/277	267/270	261/263/268	255/259	258/282	267/272
MSE0051	14	186	199	201/203	186/189/191	174/176	207	188/190
MSE0052	13	263	261/263	263/270	261	261/263	261/266	261
MSE0053	12	253	-	255/257	248/251/254	251/254	253/261	263
MSE0054	14	180/181/187/189/194	180/187/189	185/191/194/197	181/185/189/193	180/194/201	193/197/201	191/197/199
MSE0055	14	258	258	258/262	258	256/260	237/258	260
MSE0056	13	233	223	225/233	225/227/231	220	215	220/230
MSE0058	14	225/266/279/284	210	258/264/268	263/264	267/298	284	262
MSE0059	14	200	196/212	212/235	204/211/213/219	217	218/221	210/223
MSE0061	13	139	141/147	158	137/146	149/166	155	146
MSE0062	14	131	131/133	119	119	129/139	119/121	118
MSE0063	14	238/251	236/245	233	234	230/237/251	245	233/241
MSE0066	5	250	250	-	250	250	-	-
MSE0067	4	147/152	147/152	-	-	-	-	-
MSE0068	14	264	269	258/259	251/253/258/263	246/253	242	259/261
MSE0069	11	200/217	198/211	197	211/218	214	209/211	211/213
MSE0070	11	134	141	-	138	125	135	125/129/151
MSE0071	13	126/132	114/118	128/130	113/118/120/130	128	113	107/114/128

Table 2.10. continued.

	subgenus *Camellia*				subgenus *Metacamellia*		subgenus *Protocamellia*
Marker name	*C. kissi*	*C. oleifera*	*C. sasanqua*	*C. furfuracea*	*C. cuspidata*	*C. salicifolia*	*C. granthamiana*
MSE0019	125/127	129	126/137/161	129/161	136	108/147	124/129/156
MSE0021	-	-	-	278	309	271/278	-
MSE0022	178/186	186/190/194	159/176/182	192	176/196	184/194	157/159/178
MSE0023	195	180/184/192/203	184/186/198/211/219	191/195	192	-	198
MSE0024	265/272	266/273/291	245/269/275/287	261/282	253/255	251	254/260/267
MSE0025	284	287/295/306	281/288	266/284	291	302	275/281/284
MSE0026	289	282/289/296	284/285/292	291	285	282/284	272/276
MSE0027	-	-	-	114	124	-	-
MSE0029	-	-	338	367	388/390	374	368/374/390
MSE0030	255	244/261	241/250/252/255/259	261/274	257	252/257	254/255
MSE0035	227/231/235/239	222/233/235	226/234	212/217	-	229/248	222/227/235
MSE0037	164/231/245	168/218/221/222	218/227	256/262	208/214/228	228	167/169/185/222/227/23
MSE0038	313	306/308/314	307	307/308	303/305	309/324	299
MSE0039	154	151/157	146/152/154/162	156/177	135/136/144/146	154	148
MSE0040	142/147	142/149/151/153/160	153/155/161/168/170	151/153	144/147	140	155/160/164
MSE0042	107/110	107/112	102/112/114	-	-	106/110	101/110
MSE0043	164/191	162/168/172	184/191/194	-	177	194	167/169/184
MSE0044	116/127	127/129/131	121/131	132	-	-	130/138/144
MSE0045	236/242	228/236/240	231/232/235/237/241	218	236	234/238	229/234
MSE0047	280	268/280	264/277/280/285	268/278	253/256	245	284/289
MSE0049	211	211	209/211/215	213/217	211	224/226	228/230/261
MSE0050	256/265	255/263/267	262/267	255/262	272/277	268/271	267
MSE0051	185/191	185/191/199	188/199	184/190	195	191/201	186/192/201
MSE0052	261/264	261	261/263	261	-	261	261
MSE0053	249/264	250/259/264/275	246	251	-	250	270/283
MSE0054	182/193	182/189/194	180/189/191	187/191	191/194/197	194/197/202	189/197/199
MSE0055	244/258	258/260	258/260	258	258	258	258/259
MSE0056	-	235/239	229/235/241/253/256	233	217	229	220
MSE0058	275/281	222/262/273/288/302	260/267/279	230	212	189	213/218
MSE0059	220/223	194/198/212/213	194/196/198/203	211	201	220/228	198/203/210
MSE0061	167	148/162	141/142/147/149/157	-	157/161	142/151	137
MSE0062	123	131	114/118/119/135	135/139	123/131	129	119/121
MSE0063	247/251	230/234/245/253/256/26	233/236/256	230/234/237/241/243	220/230/253/262	234/238/245	234/247/249/251
MSE0066	-	-	-	-	-	-	250
MSE0067	-	158/166	149	-	-	-	-
MSE0068	253/264	249	249/253	257/258	249	264	251/264/266
MSE0069	207	221	212/237	198	-	-	-
MSE0070	125	134/141/146	134	-	142	-	129/135
MSE0071	114/120	117	118	118/120	-	114/117/126/145	118

Table 2.10. continued.

Marker name	*Number oftransferable species*	subgenus *Thea*		subgenus *Camellia*				
		C. irrawadiensis	*C. taliensis*	*C. japonica*	*C. pitardi*	*C. hongkongensis*	*C. chekiangoleosa*	*C. saluenensis*
MSE0072	14	298/320	296/300	314/319	295/308	279/294	301	304
MSE0074	13	212/218	210	204	204/206/210	206	204	206
MSE0075	14	166/175	175/182	152/164/166	166/171/177	166/185	156	166
MSE0076	13	247	247	247	247/249	247/249	247	245/247
MSE0077	14	154	128	117/128	123/132/135/136/147	119	136/151	138
MSE0078	14	144/150	148/166	144	139/156	144	141	148
MSE0079	14	103/105/112	103/111	101/107	95/97	105	103/105	95/97
MSE0080	13	238/245	238	238	240	238	238	-
MSE0081	8	118	118/126	-	116/120/126	-	126	116
MSE0082	13	159/162/165	151/159/166	164	149/162/163	151/163	151/162	150
MSE0083	14	237	253/256	249/259	242	254/257	240/246	247
MSE0084	14	398/416	408/410	404	403/404	422	404	404
MSE0087	14	280/282	279/280	282	278/280	273	282/286	278
MSE0088	3	189	180	-	-	-	-	-
MSE0089	14	292/303	292	285	286/288/290	297/305	295/301	290
MSE0094	14	195	193	184	186	190	184	188
MSE0096	14	228	241/243	243	231/236/241	248/261	233/235	225/234/239
MSE0098	14	257/275	266	266	260/263	263	260	263
MSE0099	14	299/301	301	302/308/314/316	299/302/306/331	306	302/325	299/314/317
MSE0100	7	254/262	255/258	-	260/264	252/260	-	267
MSE0101	14	247/263	247/250	263	240	252/255	263	240/250
MSE0102	12	331	320/326	323	321/327	320/328	321/335	-
MSE0103	13	174/182	176	190	170	190	-	170
MSE0106	14	150/165	152/165	171	151	175/179	161/179	163
MSE0107	14	287/289	311	310/311	313/319/321/325/329	306	306/332	315/327
MSE0108	14	256	266	249/252	250/252/256	252/257	253/256	255
MSE0109	5	-	-	117	109	-	-	-
MSE0112	14	283	283	285	283/285	283	285	283
MSE0113	0	-	-	-	-	-	-	-
MSE0114	2	199	202	-	-	-	-	-
MSE0116	14	179/183	179	178/188	167/171/173/179	168	177/183	167/171
MSE0117	14	110	108/110	114/117	114	114/116	116	114

Table 2.10. continued.

	subgenus *Camellia*				subgenus *Metacamellia*		subgenus *Protocamellia*
Marker name	*C. kissi*	*C. oleifera*	*C. sasanqua*	*C. furfuracea*	*C. cuspidata*	*C. salicifolia*	*C. granthamiana*
MSE0072	308/312	307/315	313/318	299	294	292	305/316/318
MSE0074	204/206	202/206	204	203	-	204	204/223
MSE0075	173/175	166/175/182	160/175	173/187/199	?/179/187	170/175	174/189
MSE0076	254	247/249	-	247	247	247	247
MSE0077	128	128/138	117/128/132	132/138	128/129	124/134	134/142/145
MSE0078	146	146/150	139/148/160/166/168	158	146/148	141	148/156
MSE0079	99/101	97/111	97/99/101	103	107/116	105/112	95/111
MSE0080	238	227/234/238	238	236/240	227	227	238
MSE0081	-	-	-	104/106	105/108	-	116/128
MSE0082	164/165	154	154/156/164	-	160/161	174/176	162/163/166
MSE0083	240	240/244/247	236/244	256/268	243	249/252	247
MSE0084	403/416	404/409	404/409	414/429	412	412/424	404
MSE0087	275	283	278/286	272/273	273/276	273/278	272/278
MSE0088	-	-	-	-	-	195/196	-
MSE0089	295/305	285/288/295/299/303/30	290/293/295/297/301	294/311	295/307	283/285	288/290/292/307
MSE0094	184	184	184	188	186/188/191	190/191	186/188/227
MSE0096	237/251	229/244/258	229/247	235	222/230/240	234	236/250
MSE0098	260/269	263	254/263/269	263/266	263	257/266	263/266
MSE0099	306/312/323/329	297/307/308/321/325	308/317	306	299	303	299
MSE0100	-	-	-	-	247/250	-	238/245
MSE0101	263	240/263/269	240/263	250	244	240/252	240/247
MSE0102	325	312/315/324/328	300/315/318/346	311	309/345	336/372	-
MSE0103	182/194	184	190/198	182/184	182/196	186/188	174/182/190
MSE0106	152/155	144/151/167/183	152	165	163	149	159/175
MSE0107	317/321	315/319/323	311/315/319/323	308/317	304/317	289/303	308/319/321/323
MSE0108	249/255	251/256	249/251	252	247	253/256	251/252
MSE0109	-	117/125/127	125/127	-	-	-	103
MSE0112	285	285/289	285	283	285	283	285
MSE0113	-	-	-	-	-	-	-
MSE0114	-	-	-	-	-	-	-
MSE0116	174	169/172/173	174/177/185	166	165/175	172	162/172
MSE0117	112	114	106/116/117	114	114	125/131	114

Table 2.11 Polymorphism information of genomic SSRs.

Marker	Size range	No. alleles	H_E[a]	H_O[b]	PIC[c]
MSG0004	251–258	2	0.38	0.50	0.30
MSG0005	179–194	6	0.81	0.63	0.78
MSG0006	217–249	7	0.83	0.31	0.81
MSG0007	155–186	9	0.81	0.56	0.79
MSG0008	270–291	6	0.61	0.25	0.58
MSG0009	151–166	3	−	−	−
MSG0012	275–284	4	0.56	0.38	0.49
MSG0013	227–241	3	0.56	0.93	0.46
MSG0015	357–410	12	0.90	0.46	0.89
MSG0017	215–243	11	−	−	−
MSG0018	200–233	10	−	−	−
MSG0023	277–317	13	0.89	0.81	0.88
MSG0024	219–255	12	0.88	0.88	0.87
MSG0025	175–212	10	0.81	0.50	0.79
MSG0027	225–284	6	0.70	0.63	0.66
MSG0029	88–109	7	0.69	0.73	0.66
MSG0030	288–361	5	−	−	−
Mean	–	7.5	0.72	0.58	0.69

[a] H_E, expected heterozygosity.

[b] H_O, observed heterozygosity.

[c] PIC, polymorphism information content.

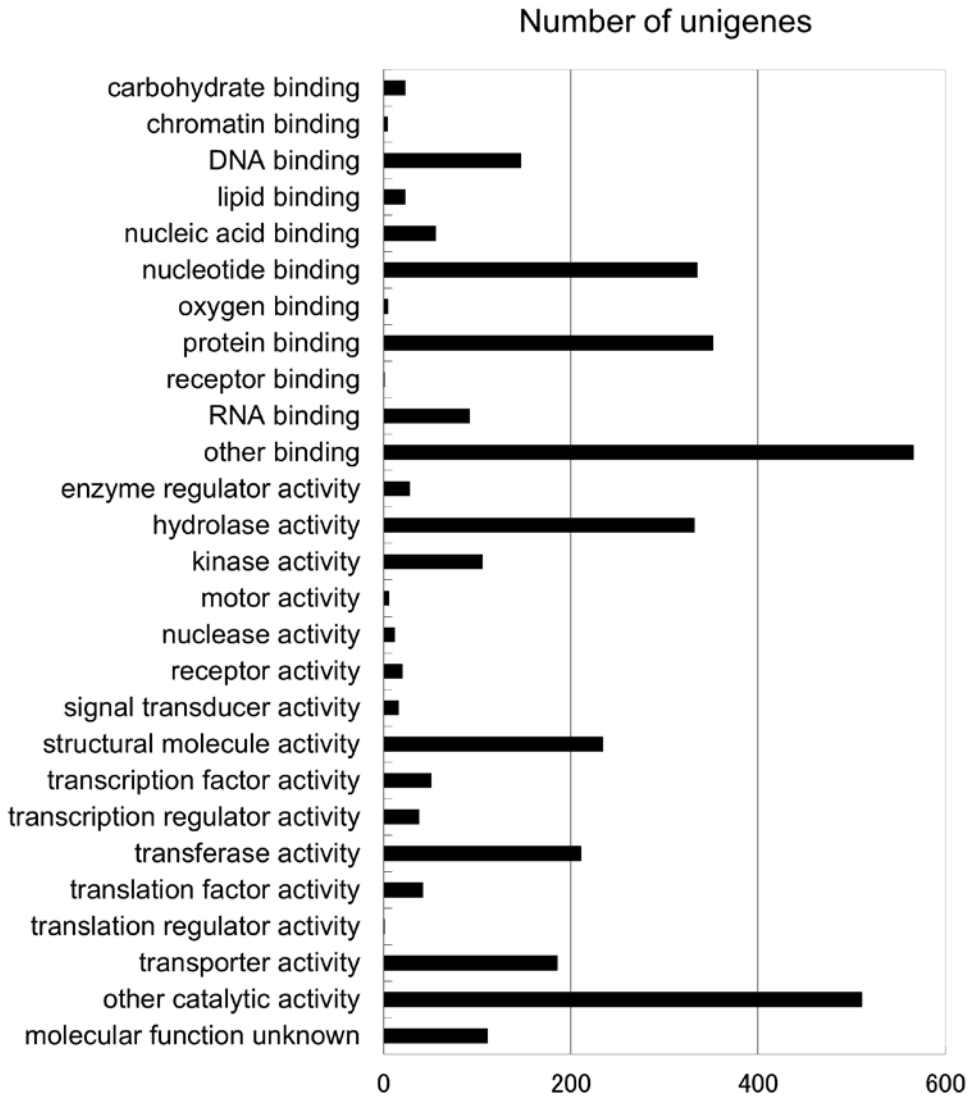

Fig. 2.1. GO Slim term annotation of tea unigenes in the 5.3-k set. Each unigene was assigned one or more GO Slim terms by the Blast2GO software.

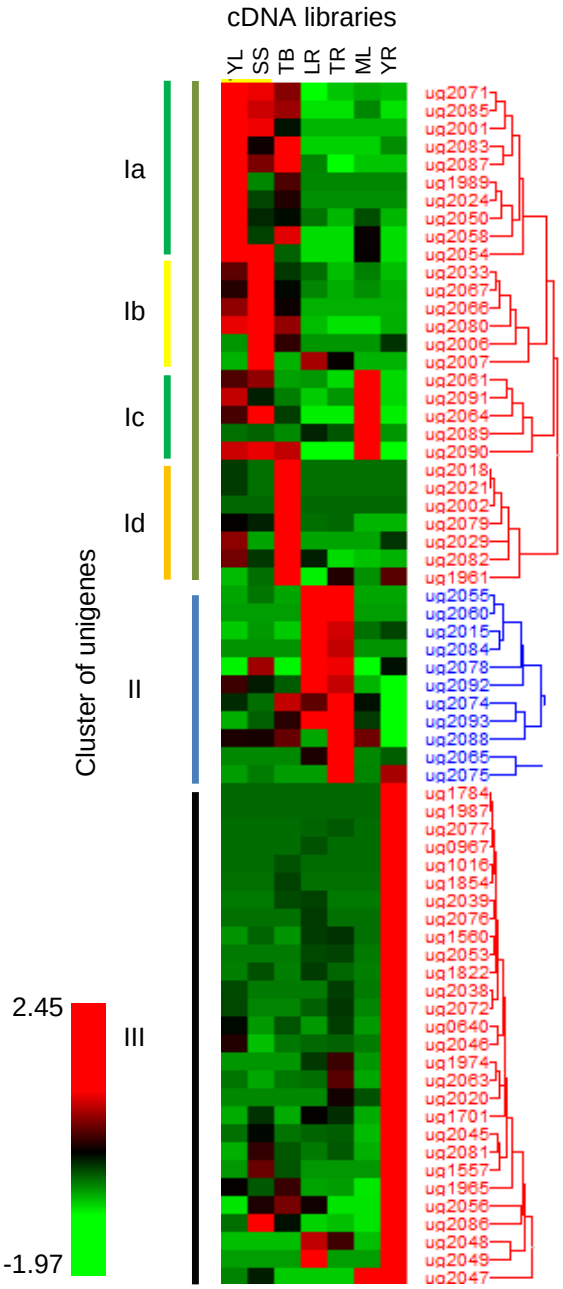

Fig. 2.2 Hierarchical clustering of 67 unigenes showed differential expression among the seven cDNA libraries (YL, young leaves; SS, shoot stems; TB, terminal buds; LR, lateral roots; TR, tap roots; ML, mature leaves; YR, young roots). The unigenes were grouped into three major clusters (indicated by vertical color bars), and cluster I was further subdivided into four subclusters. Red cells represent normalized expression values greater than the mean for that gene; green cells represent lower expression than the mean.

CHAPTER 3 Construction of a high-density reference linkage map of tea

3.1 Introduction

Genetic linkage maps are important resources for genetic and breeding studies. A linkage map for tea would allow the genetic analysis of important traits through the development of selective markers for use in MAS, and would supply DNA markers for the investigation of genetic diversity of germplasms. This chapter considers the role of MAS in tea breeding.

Traditionally, tea is grown from seed, but the resultant plants do not have synchronous bud break, uniform harvest times, or uniform quality. In contrast, clonal cultivars, now grown in many countries, have those three features, making them well suited to machine harvesting. To breed clonal cultivars, breeders have to select a large number of individuals with desirable characteristics from a large number of F_1 segregating populations and propagate them clonally for evaluation. Elite clones are released as cultivars. The production of high-quality tea leaves is the most important breeding objective, but evaluation is laborious and time-consuming. In addition to being a lengthy process, tea breeding also requires large fields. For these reasons, it is more efficient and effective to multiply early-stage populations and select breeding materials by the evaluation of DNA markers for important agronomic characters that are controlled by a few loci before the seedlings are planted in the field. At Makurazaki Tea Research Station, a breeding system that combines MAS in the juvenile phase with the acceleration of generations has shown successful results (Tanaka 2006). It would be preferable to perform MAS for many characteristics in many different populations, so a method of genetic analysis is needed for a wide range of genetic materials. This requires a reference map with many landmark markers.

When genetic diversity or phylogenetic relationships are investigated, it is preferable to select DNA markers for genotyping evenly across the whole genome, not from a specific region, because the regions of domestication genes could have been subjected to artificial selection during the history of the crop, or because there could be selective sweeps in one or more regions related to adaptation to the cultivation

environment resulting in low genetic diversity in specific regions. By contrast, a reference linkage map consisting of many highly polymorphic DNA markers distributed genome-wide is a better resource for the selection of DNA markers for genetic diversity research.

A few linkage maps of tea based on the pseudo-testcross theory (Grattapaglia and Sederoff 1994) have been reported (Hackett et al. 2000, Ota and Tanaka 1999, Tanaka et al. 1995). These maps were constructed through the use of dominant markers such as RAPD or AFLP markers. By using pseudo-testcross theory, it is easy to construct accurate linkage maps for small F_1 populations derived from heterozygous parents, because the segregation ratio of each marker is 1:1 and the linkage relationships among them are simple and clear. However, the utility of dominant markers depends on the particular material being tested, as markers are not universal. In addition, the relationships among the linkage groups (LGs) in the previously reported maps are unknown. In a few fruit trees, reference linkage maps have been developed by using co-dominant CAPS and SSR markers (Omura et al. 2000; Yamamoto et al. 2002, 2005, 2007). Here, I developed a high-density reference map, using both newly developed SSR markers in this study (Chapter 2) and some previously reported SSR and CAPS markers (Ogino et al. 2009, Ujihara et al. 2011) as landmark markers, and RAPD markers, which were not used as landmark markers because of their dominance.

3.2 Materials and methods

3.2.1 Plant materials

An F_1 segregating population consisting of 54 clones was derived from reciprocal crosses between the Japanese green tea cultivar ‘Sayamakaori’ and the line Kana-Ck17. ‘Sayamakaori’ was selected from a population derived from naturally pollinated seedlings of ‘Yabukita’ (Fuchinoue et al. 1972); its pollen parent is believed to be germplasm from China (Tanaka et al. 2001). Kana-Ck17 was introduced from China. Chinese tea materials have greater genetic diversity than Japanese materials (Kaundun and Matsumoto 2003, Wachira et al. 2001, Yamaguchi and Tanaka 1999). I expected there to be high heterozygosity between the genomes because of the lack of duplication in the pedigrees of the parents and because tea is an out-crosser.

3.2.2 Development of SSR markers

SSR markers with names beginning with ‘TMS-LA’ and ‘MSG’ (Table 3.1) were developed from microsatellite-enriched libraries. These libraries, enriched for GA and GT, were developed by Genetic Identification Services Inc. (Chatsworth, CA, USA) from DNA of the line KM62, extracted by using a modified CTAB method (Tanaka et al. 2001). Markers were sequenced on a model 310, 3130*xl*, or 3700 Genetic Analyzer (Applied Biosystems). The primer sets were designed with Primer Express (Applied Biosystems) or Read2Marker (Fukuoka et al. 2005) software. SSR markers with names beginning with ‘MSE’ (Table 3.1) were derived from EST databases; the corresponding primer sets were also designed with Read2Marker. Marker TMS3 was developed by using a PCR method (Tanaka and Ikeda 2000).

3.2.3 SSR marker detection

DNA was extracted from fresh leaves by a method using diatomaceous earth and a spin filter (Tanaka and Ikeda 2002). The PCR conditions were similar to those previously described (Tanaka et al. 2003). PCR reactions (10 µL total volume) contained 0.5 units of AmpliTaq DNA Polymerase (Applied Biosystems), the attached reaction buffer, 2.0 mM $MgCl_2$, 0.16 mM each dNTP, 1% formamide, 2 ng/µL template DNA, and 2 µM of each primer DNA. The PCR temperature conditions were based on the touchdown PCR technique (Don et al. 1991). The PCR program consisted of an initial 94 °C for 5 min; 34 cycles of 94 °C for 30 s, a decreasing annealing temperature (described next) for 60 s, and 72 °C for 30 s; and a final extension at 72 °C for 10 min to allow complete double-strand DNA synthesis. The annealing temperature, which was 62 °C in cycle 1, was reduced by 0.5 °C per cycle during cycles 2 to 14 and maintained at 55 °C for the last 20 cycles. PCR products were post-PCR labeled (Kukita and Hayashi 2002), then separated and detected on a model 310, 3130*xl*, or 3700 Genetic Analyzer. The sizes of the amplified fragments were calculated against internal-standard DNA (Gene Scan 500 LIZ, Applied Biosystems) in the GeneMapper software (Applied Biosystems).

3.2.4 CAPS marker detection

Total DNA was extracted by using a modified CTAB method (Kaundun and Park 2002). PCR was carried out in a 25-µL solution containing 1 µL of template DNA, 0.625 µM each primer, 2 µL of each dNTP, 2 mM $MgCl_2$, 1× ExTaq buffer, and 0.625 U of ExTaq DNA Polymerase (Takara, Japan). The PCR program consisted of an initial 94 °C for 5 min; 35 cycles of 94 °C for 30 s, 60 °C for 1 min, and 72 °C for 90 s; and a final extension at 72 °C for 10 min. Two microliter of the PCR reaction mixture was digested by exact restriction endonuclease (Table 3.2). After incubation for 2 h at 37 °C, the digests were separated by 2% TAE–agarose gel electrophoresis; the agarose gel and the TAE buffer both contained 0.5 µg/mL ethidium bromide. CAPS primer sequences are shown in Table 3.2.

3.2.5 RAPD marker detection

Template DNA for RAPD marker detection was prepared as for SSR analysis. PCR reactions (10 µL total volume) contained 0.5 units of AmpliTaq DNA Polymerase (Applied Biosystems), the attached reaction buffer, 2.5 mM $MgCl_2$, 0.125 mM each dNTP, 2 ng/µL template DNA, and 2 µM primer DNA. The PCR program consisted of an initial 93 °C for 5 min; 40 cycles of 93 °C for 1 min, 42 °C for 1.5 min, and 72 °C for 1 min; and a final extension at 72 °C for 10 min to allow complete double-strand DNA synthesis. The PCR products were separated by electrophoresis in 2% agarose gel, stained with ethidium bromide, and visualized under a UV transilluminator. The name of each RAPD marker refers to the primer name and the expected band size; for example, ‘OpA_12_600’ represents the marker band of 600 bp obtained with the primer OpA-12 (‘Op’ means Operon Technologies, Inc., Alameda, CA, USA).

3.2.6 Mapping algorithm

The F_1 segregating population was not very large, so Ideveloped a three-line reference map by using information from the parents as well (Fig. 3.1). This reference map was designed to provide an accurate marker order and linkage distances between markers despite the limited population size. AntMap (Iwata and Ninomiya 2006) software, which is based on ant-colony optimization theory (Dorigo et al. 1996), was

used to construct the linkage maps. On the basis of double pseudo-testcross theory (Grattapaglia and Sederoff 1994), the F_1 segregating population was regarded as resembling a first-generation backcross (BC_1) population for the purpose of linkage analysis. The segregation data were first sorted into data sets derived from each parent (Fig. 3.1), then a map of each parent was constructed from the sorted data sets. As SSR marker bands generated by the same primer sets are likely to amplify the same locus in each parent, SSR markers tend to map within LGs in the same order in both parental maps. Thus, the amount of genotype data was doubled for those markers that detected polymorphisms within the two parents. I manually selected those common markers, which I assumed to be derivatives from the same locus in both parents. This selected marker set was designated as the core marker set. I then analyzed the linkage relationships within the core marker set on the basis of the combined (54 + 54 = 108) genotyping data, and constructed a new map that consisted of only the core marker set. This merged map was designated as the core map. To present the relationship between the LGs of the core map and the two parental maps, I used the MapChart software (Voorrips 2002) to sandwich the core map for each LG between the parental maps.

3.2.7 RAPD STS

Methylene-blue-stained agarose gel blocks containing target RAPD bands were cut out, and the DNA was re-amplified by PCR for 15 cycles, as described above for the detection of RAPD markers. Re-amplified PCR products were ligated into the pGEM-T vector (Promega, Madison, WI, USA), which was transformed into *E. coli* strain DH5α competent cells for multiplication and sequenced with a 3130*xl* Genetic Analyzer.

To clone the sequences that flanked the RAPD STS markers, I performed thermal asymmetric interlaced (TAIL) PCR (Liu et al. 1995) using arbitrary primers MAT01 (CNSCTSCTNTWTCTT), MAT02 (CNNCWNCAANAWCAA), MAT03 (GNSGASGANAWAGAA), and MAT04 (GNNGWNGTTNTWGTT), and specific primers designed from the STS sequence information. PCR solutions (10 µL total volume) were the same as those used for SSR marker detection, with an arbitrary primer concentration of 50 µM. The temperature program followed the TAIL-PCR protocol (Liu et al. 1995), integrating the touchdown PCR technique. The annealing temperature

for the first PCR, which was 60 °C during the first cycle, was reduced by 1 °C per cycle during the first 5 cycles. In the asymmetric super-cycle, the temperature for non-specific annealing was held at 38 °C, and the temperature for specific annealing started at 60 °C and was reduced by 0.4 °C per cycle for 15 super-cycles. For the second and third PCRs, the temperature for non-specific annealing was again held at 38 °C, and the temperature for specific annealing started at 60 °C and was reduced by 0.5 °C per cycle for 12 super-cycles. For the third PCR, primers with single nucleotides (A, T, G, or C) added to the 3′-end of the arbitrary primer sequences were used for selective amplification of the flanking sequence.

3.3 Results

I constructed three-line reference linkage maps of ‘Sayamakaori’ and Kana-Ck17 using SSR, RAPD, and CAPS markers (Fig. 3.2). The markers on the maps are described in Tables 3.1 and 3.2. The map of ‘Sayamakaori’ consisted of 701 loci, 363 of which were SSR markers. It encompassed a genetic distance of 1305 cM, with an average distance between markers of 1.86 cM. The map of Kana-Ck17 also consisted of 701 loci, 354 of which were SSR markers. It encompassed 1298 cM, with an average distance of 1.85 cM. Both maps contained 15 LGs (Table 3.3). All 441 SSR markers with normal segregation within the F_1 population could be mapped to one of the 15 linkage groups. These two maps each included the 279 SSR markers and 2 CAPS markers in the core marker set.

The preliminary core map had a large gap in LG 5. To cover this gap, I cloned and sequenced RAPD marker ‘OpT_18_200’ from the map of Kana-Ck17. Because there is no polymorphism for this sequence in ‘Sayamakaori’, I amplified the flanking sequence by TAIL-PCR and sequenced it. This revealed 2 SNPs with heterozygosity in ‘Sayamakaori’ in the flanking sequence (accession numbers AB623060, Sayamakaori; AB623061, Kana-Ck17). These SNPs were mapped to the same locus as RAPD marker ‘OpM_11_750’ on the ‘Sayamakaori’ map. The segregation patterns obtained for ‘OpM_11_750’ and the new marker were identical, so I mapped the new STS marker (designated ‘OpM_18_200STS’) on the core map.

The revised core map contained 15 LGs with a total length of 1218 cM without

gaps. The average distance between markers was 4.35 cM.

3.4 Discussion

3.4.1 Characterization of the new linkage maps

The core map and the parental maps each covered 15 LGs; this corresponds to the basic chromosome number (n = 15) of tea. All 441 SSR markers with normal segregation within this population could be mapped to one of the 15 LGs. The total length of the core map is 1218 cM, and that of the combined (three-line) map is 1317 cM. The genome size of tea is estimated to be 4 Gbp (Tanaka et al. 2006), so 10 cM of genetic distance corresponds to 30 Mbp of physical distance on average.

3.4.2 Illustration and application of the reference linkage map

In the reference map, the core map for each LG is sandwiched between the parental maps. Although obtained from the same F_1 population, the sets of linkage information for each parent are independent. If the maps of the two parents were merged into a single map by using the landmark markers, the order of the non-landmark markers would be inaccurate. This three-line illustration presents the marker order reasonably accurately.

It is common for marker order in linkage maps to be inaccurate because of missing data, especially when the population size is small. I used automated DNA sequencers to detect the size of the amplified fragments from the SSR marker analysis. In this method, some background peaks hampered the detection of particular alleles; exampes include MSG0213 in LG 4 of ‘Sayamakaori’ and MSE0230a in LG 3 of Kana-Ck17. For such markers, the core map is likely to be more accurate than either parental map, because it is based on double the amount of linkage data per marker.

In addition, the core markers are likely to be more informative and more useful than other markers for genetic analysis in other populations. For a marker to be designated as a core marker, both parental materials had to be heterozygous for the marker, and the segregation of the alleles derived from one parent had to be detectable without the interference of alleles from the other parent. That is, if the genotype of one parent was AB, the genotype of the other parent had to be AC, BC, or CD. In other

words, a core marker must have had at least three distinguishable alleles within the two parental lines. To select markers for genetic analysis, it is efficient to choose markers from the core marker set first, followed by co-dominant markers on the parental maps. If there are no polymorphic markers within a target area, it is possible to search SNP markers from the sequence information of SSR markers. Even if no SSR markers are found in the target region, it is possible to search SNPs in the flanking regions of SSR markers, or even from the sequence of a RAPD band or its flanking region, as I did in the case of the OpT_18_200STS marker.

For genetic diversity studies, DNA markers on the core map which have at least three alleles per locus are a good resource for the analysis of genetic diversity by scanning the whole genome.

The linkage map presented here should be useful as a reference map that will play an important role in establishing the alignment of genomic sequences of tea as I work toward full-genome sequencing. The development and mapping of additional markers will further advance the accuracy and density of the available maps.

Table 3.1. SSR markers on the reference map of tea

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSE0158	LG01	C.M.S.	tcaaaaatgctttcgaatccgc	cttgacgaagtgttcgtaggcttg	(tc)3, (cac)3, (cca)7	7	324	AB623897	
MSE0039	LG01	C.M.S.	ccattctgctgcaattacacatcc	gattttggtaagcggctcattgtc	(ag)21	21	171	AB485968	Ogino *et. al.* (2009)
MSG0361	LG01	C.M.S.	agatggaggtagagagagaggcag	gtttgtccctctcattttcaacgc	(ag)26	26	291	MSG0361	
MSG0258	LG01	C.M.S.	actcatcaccatgccttctccatc	gtttagctcaactggtggaacctcaact	(tc)25, (ta)4, (ta)3	32	289	MSG0258	
MSE0269	LG01	C.M.S.	ttgacggcatttttgtgtgatagg	tttcctatcatttcccacacgctt	(tc)9, (ca)3	9	241	AB623945	
MSE0316	LG01	C.M.S.	ggaaagagggaaccaccaaagaat	ccattcagatagccaaacaaagcc	(tc)14	14	313	AB623970	
MSE0282	LG01	C.M.S.	tacggaaggaaggaaggaaggaag	ttttgatgcgtgagggttaaaggt	(cca)7, (ccg)3, (tca)3	7	322	AB623950	
MSG0121	LG01	C.M.S.	gagcaaagaagttgtgatttcgcc	attttgagatttcgaagctccacc	(ag)11	11	264	AB623989	
MSG0835	LG01	C.M.S.	atatttgggtggtgattgaggatt	gtttcaccaattaatccaccccaacaaa	(at)3, (ag)12	15	239	AB624255	
MSG0542	LG01	C.M.S.	atatcacataggtttgtcccacca	gtttagagggtaaggcggagtggaatag	(tc)12, (tc)9	21	287	AB624125	
MSG0811	LG01	C.M.S.	acaccacaccacaccacactttct	gtttggtctgaagctccaaagtgaa	(tc)15	15	151	AB624246	
MSG0814	LG01	C.M.S.	agaaagctcaaacatggcgaaaag	gtttggaccgaccactcaagcagtctat	(ag)15	15	114	AB624248	
MSG0318	LG01	C.M.S.	atgtgttcgccattgctagggatt	gtttagcacgctccatttttcagagc	(ag)3, (ta)5, (ac)3, (ag)17	28	300	MSG0318	
MSG0274	LG01	C.M.S.	agttcaccaacgatctctctctcg	gtttgacagttcgatacctgtcagacaatca	(tc)4, (tc)11, (ac)11, (ta)5	31	273	MSG0274	
MSG0188	LG01	C.M.S.	atgggtttgagattagggtttcaa	gtttggtaatatgttccccagtcccg	(tg)12, (ag)17, (tg)3, (tg)12, (cg)4, (gt)3, (ag)3, (ag)4, (ag)6, (ta)4, (taa)3	71	299	MSG0188	
MSE0103	LG01	C.M.S.	gcaagcaaaggcattacaaaaagc	ttcaccaactgaaagaatccccat	(tc)14	14	183	FS951354	Chapter 2
MSG0531	LG01	C.M.S.	agaagaatagtcggagttgcagcg	gtttcaacttgcaaaaaccaactcag	(tc)13, (ta)8	21	262	AB624115	
MSG0740	LG01	C.M.S.	atagaacaccacctgactccaatg	gtttgttttgtctggacgtggctaa	(tc)12, (tc)5	17	180	AB624217	
MSE0009	LG01	SAK	cgctcttcttaagtggaaaggggt	tggatacgcaaatctccaaacaga	(ct)3, (tc)15	15	241	CV014223S	
MSG0410	LG01	SAK	attgataatgaaatccgtcggtgg	gtttacgggcttatttgtttggaaccat	(ggt)3, (ag)22	25	172	MSG0410	
MSE0034	LG01	SAK	cgctctctgttccgaagcttttag	ccttcacctccaaacaaccaaaac	(tc)14, (tc)4, (tc)4	14	192	FS951853	Chapter 2
MSG0726	LG01	Kana-Ck17	agagagcttgctcatcaaagtagtgg	gtttacctccaatttgtgaaatgccaac	(ag)17	17	103	AB624212	
MSG0793	LG01	Kana-Ck17	atatcgagtccgaactgtagcgag	gtttgggtagtttacgatttgaagggaga	(ag)16	16	160	AB624235	
MSE0323	LG01	Kana-Ck17	aattcagtggcatccgagaagaag	tcacagacttactacatgaacaaggca	(ag)3, (ta)9, (tg)4, (cg)4, (gc)3	9	331	AB623972	
MSE0325a	LG01	Kana-Ck17	ggaggttcaactccagtttctcca	tttgacgtcactgcgaaatctcat	(at)3, (tc)8, (tg)3, (ta)3, (ca)3	8	324	AB623974	
MSG0448	LG01	Kana-Ck17	atttgttgccgtcgacactagaga	gtttccaccacacacacacacatacaca	(ag)3, (ag)5, (ag)8, (ag)5, (gtg)3	24	246	AB624077	
MSG0199	LG01	Kana-Ck17	aggaattttcaggtgagggaaagc	gtttggcattggtatcgtgttatttgc	(ag)13, (ag)4, (ag)4, (agt)3, (ag)9, (tg)8, (caa)3	44	280	MSG0199	
MSG0053	LG01	Kana-Ck17	agacaacggagcactcatgtcatc	gtttcaggtttcgaactgttgtgtgcat	(ag)14	14	173	AB623984	
MSE0275	LG01	Kana-Ck17	aactctctgcccaacacgttacca	cttaggtttcgaagacagcccaga	(tc)4, (tc)5, (cct)4	5	216	AB623947	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0476	LG01	Kana-Ck17	atggagggtcgcctatgatggagt	gtttggagggagacttttgatttttggc	(ga)3, (ag)14, (ag)6	23	290	AB624093	
MSE0177	LG02	C.M.S.	acccatcaccacttcctttccttc	ggtccatcaccgacttaacgagag	(tc)12	12	245	AB623906	
MSG0460a	LG02	C.M.S.	agatgaaggggacaacccaatttt	gtttcaatggttgagattgtttgctcac	(ag)20, (gaa)3	23	151	AB624084	
MSG0613	LG02	C.M.S.	acatacacgtattcacaagttggtcc	gtttccaaactaatgcgatccatacca	(tc)16, (ta)3	19	246	AB624156	
MSE0207	LG02	C.M.S.	ttacgagtgtgcggagtctacgag	tccactgagaagacgatgcgttta	(ag)14	14	254	AB623916	
MSG0471	LG02	C.M.S.	atccctcatttggactttgcattt	gtttgcagggttagcatttttggaac	(ag)17, (ag)6	23	187	AB624090	
MSG0544	LG02	C.M.S.	atatgaagtgggtggagggtttct	gtttcagtcgagctcaattgttcg	(ag)3, (ag)18	21	202	AB624126	
MSG0747	LG02	C.M.S.	atctgtctctgccgatctctgttg	gtttaaagagaagaagcatttgggaggg	(tc)17	17	198	AB624220	
MSE0173	LG02	C.M.S.	gtgttcaccaacaactcaccaagg	tgtcgaacaaagatacaccccaaa	(tc)10, (ta)9	10	285	AB623905	
MSG0429	LG02	C.M.S.	aggaccgttcttccctacctgtaa	gtttgagattgaggatgtggtcgttgt	(tc)24	24	298	AB624066	
MSE0280	LG02	C.M.S.	gtcctctctctttcatcgcggag	atgaatcgacgacgacgaaaacac	(tc)4, (tc)8, (ag)3, (ta)3, (ta)6	8	207	AB623949	
MSG0430	LG02	C.M.S.	aggagaggagatgaagagttgctg	gtttaagcaaacccaagaggcgtatttt	(ag)4, (ag)4, (ag)16	24	215	AB624067	
MSG0655	LG02	C.M.S.	atggtatcagtggattgggaaagg	gtttctctgggttcttcccaagttgtc	(ga)3, (ag)13, (ga)3	19	226	AB624175	
MSE0263	LG02	C.M.S.	agtctcagcatcaccgattcacag	tgggtctgagttggttgtgattgt	(ag)11	11	332	AB623942	
MSG0609	LG02	C.M.S.	acacccaagtccaaactcaaactc	gtttcgactgatcggtgaccttctcc	(ag)19	19	158	AB624153	
MSG0518	LG02	C.M.S.	atgtgagtgagattaagcgagggg	gtttactttcaggatacggtagggcga	(ag)3, (ag)19	22	296	AB624107	
MSE0023	LG02	C.M.S.	tgaaacaaccacctcccttgatct	agcagcgaagaaggattcattacg	(tc)13, (tc)7, (ct)3, (ta)5	13	213	AB485966	Ogino *et. al.* (2009)
MSE0305	LG02	C.M.S.	tggggtcaggttcagaaataatgg	tcctaccaaaaggcaaaaagagaca	(tc)3, (tc)18	18	271	AB623965	
MSG0694	LG02	Kana-Ck17	atcctcacctccattaaggagcaa	gtttaaaggccagaagtgtagaactgcg	(tc)18	18	167	AB624198	
MSG0394	LG02	Kana-Ck17	aggatgtcggagggagatttagag	gtttccctaaaattcattctcacgcct	(ga)3, (agg)3, (ac)3, (ag)16	25	283	MSG0394	
MSE0297	LG02	Kana-Ck17	agctgagtcctagagacccccact	tcagtggattgattgagattggga	(tc)6, (ta)3, (ta)3, (at)3	6	260	AB623959	
MSG0590	LG02	Kana-Ck17	atacattggtgctcattcccaagc	gtttgacactcgtcgctaggctactg	(ta)4, (ag)16	20	116	AB624143	
MSG0096	LG02	Kana-Ck17	taaatccccgaaatccctaaatcg	tttgttccttcggagtatgatcgg	(tc)17, (ac)10, (at)3	17	234	AB623986	
MSE0223	LG03	C.M.S.	tttgtgctctcatcaccatcatca	attctcaccttggtgcctctcttg	(caa)12	12	269	AB623926	
MSE0231	LG03	C.M.S.	gaaagagagattgtgggttgtggg	ttggcgataacttcgagacacaaa	(tc)5, (tc)3, (tc)13	13	328	AB623930	
MSE0245	LG03	C.M.S.	ttcgttcttcggatagtctcaggc	tatcttggagagcttgttgggagc	(ag)12	12	203	AB623936	
MSG0371	LG03	C.M.S.	atggcatcgcaaaaggagaatagt	gtttgtttacattcaacaactgcccg	(tc)23, (gaa)3	26	232	MSG0371	
MSE0006	LG03	C.M.S.	ggggactcactcattcacttcaatc	atcccagctctgtaaaacgaccac	(tg)4, (ag)9, (aga)3, (ag)3	9	257	CV013588S	
MSG0219	LG03	C.M.S.	atgttgagttgtcgacccagagag	gtttacctcacctcacctgaaaaaccc	(ag)7, (ag)6, (tc)3, (tg)3, (ta)3, (ta)3, (ag)13	38	234	MSG0219	
MSG0572	LG03	C.M.S.	agcactctcaggatttctgctcgt	gtttctggtgatgaaagcccagtttc	(tc)20	20	157	AB624134	
MSG0592	LG03	C.M.S.	atagttcgatgactccgataccaa	gtttgagagagggagagcgagagaatga	(tct)3, (tc)17	20	218	AB624144	
MSG0473	LG03	C.M.S.	atgatgttgatggattggtgtgga	gtttatgtgcgggactttgtgtttttg	(ag)3, (ag)20	23	216	AB624092	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0382	LG03	C.M.S.	accaaagctgaatcattaattgcc	gtttgccatggaaggaagtgaagtgag	(ac)3, (at)3, (tc)3, (tc)16	25	281	MSG0382	
MSG0255	LG03	C.M.S.	acccttgtctattctccatcgctc	gtttctagggtttgcgttgagaccttc	(tc)21, (ac)3, (ac)3, (ac)5	32	245	MSG0255	
MSG0533	LG03	C.M.S.	agacctagccaagacaaccacacc	gtttcccctattttcccgactgtct	(ag)18, (ga)3	21	223	AB624117	
MSG0462	LG03	C.M.S.	aggacaaaccgaaacaaaatgacc	gtttgaggtcgaagtcatagcgtcatca	(ac)3, (tc)4, (tc)16	23	292	AB624085	
MSG0322	LG03	C.M.S.	attttgtgacaacactactgtgcaa	gtttgaggtcctggtggtagttccatt	(tc)12, (tc)16	28	251	MSG0322	
MSG0098	LG03	C.M.S.	caagccagagcttcctcgtatgat	tcgttcgtcaacttactcactcgc	(tg)4, (ag)11, (ag)6, (ag)4, (ag)3	11	163	AB623987	
MSE0082	LG03	C.M.S.	tacaaaatccctcccccaaaaa	tcaaactaagaaccaatgccacca	(ca)3, (ta)13	13	167	FS948901	Chapter 2
MSE0029	LG03	C.M.S.	atagccaatcaagctcctcctcct	agtctgttcctcccttgatgatcg	(ag)14, (ag)7, (gga)3	14	276	FS949897	Chapter 2
MSE0230a	LG03	C.M.S.	agtctagttgggcaacggtgagac	aaaagtcattccccttttcccctt	(tc)4, (tc)4, (tc)5	5	348	AB623929	
MSG0232	LG03	C.M.S.	atttcgatctcccagacgtgaact	gtttacagaactagaagcaccaaaccgc	(tc)3, (tc)3, (tc)3, (tc)3, (tc)24	36	214	MSG0232	
MSG0717	LG03	C.M.S.	acccctttccgatcatcaggtacta	gtttctcaccaatcaacaccgtttca	(tc)13, (ta)4	17	189	AB624208	
MSE0244	LG03	C.M.S.	tcatctccacacaacccattccta	agagagcaagtaagcctcccactg	(ag)13	13	235	AB623935	
MSE0196	LG03	C.M.S.	acacacatacaacccaacccaagc	agcctcgtataggagtgagcgaaa	(tc)3, (tc)3, (tc)9, (ac)3	9	333	AB623910	
MSG0612	LG03	C.M.S.	acagcaaagcccgatcagtaccta	gtttgagattcgctgacttagggcttca	(tc)19	19	275	AB624155	
MSE0210	LG03	SAK	tccactcttgtccggctcttctat	ttggactcatcgaaggggatctta	(ag)15	15	344	AB623919	
MSG0386	LG03	SAK	acgcatgctgatgtatgtcaagaa	gtttgggaagcacgaccgactagaataa	(tc)25	25	157	MSG0386	
MSG0666a	LG03	SAK	accctcacaacaaaccctggtcta	gtttcagtcgatttttcttccccgt	(aga)3, (ag)15	18	273	AB624182	
MSG0806	LG03	SAK	attatccgactgtgggtctttcca	gtttcttttcctgcactgtgatccgaata	(ag)16	16	163	AB624243	
MSE0045	LG03	SAK	ttggcgataacttcgagacacaaa	ctccctccctccctttaatggtt	(ag)14, (ga)3, (ag)4	14	222	AB461368	Ogino *et. al.* (2009)
MSG0727	LG03	SAK	agagcacctagagggagagtgaaa	gtttacccttcatttcctcacaacctca	(ag)3, (ag)8, (ag)6	17	257	AB624213	
MSE0061	LG03	SAK	acacacatacaacccaacccaagc	ggaggactctgcatccattacagc	(tc)3, (tc)3, (tc)9, (ac)3	9	148	FS949588	Chapter 2
MSE0077	LG03	SAK	cccaggcgtctttgtatacgtctc	tcgaaatgcactctctcattttcg	(tc)3, (tc)10, (cg)3	10	137	AB361050	Chapter 2
MSG0666c	LG03	Kana-Ck17	accctcacaacaaaccctggtcta	gtttcagtcgatttttcttccccgt	(aga)3, (ag)15	18	273	AB624182	
MSE0221	LG03	Kana-Ck17	cagcctgtactcttccattcttgc	tttttcagtcactagctctggggc	(tc)14	14	206	AB623925	
MSG0666b	LG03	Kana-Ck17	accctcacaacaaaccctggtcta	gtttcagtcgatttttcttccccgt	(aga)3, (ag)15	18	273	AB624182	
MSG0402	LG03	Kana-Ck17	atgacactcggtccataaattcaa	gtttggtgatgacacgacctagggtt	(tc)25	25	255	MSG0402	
MSG0437	LG03	Kana-Ck17	atacttttgcagacccacctcctc	gtttcccagtggcactaaagttgcctaa	(ag)17, (ag)7	24	293	AB624072	
MSE0268	LG03	Kana-Ck17	gggggttgtcagttgatttctcac	gaaaccagagactcttatggcgga	(ag)6, (ag)9, (ag)3, (ag)3, (ag)4, (ga)3	9	209	AB623944	
MSE0206	LG03	Kana-Ck17	tttcttttcggatgatcaagctcc	atatcattcccgcatccttcagag	(ag)7, (ggt)6	7	329	AB623915	
MSE0254	LG03	Kana-Ck17	atcctgtcaaagctcctcttgtgg	gagtagcattttggcttattggatcac	(ct)3, (tc)19, (at)3	19	239	AB623939	
MSG0383	LG03	Kana-Ck17	accaggtcacaagactcagaacaa	gtttcgacttcgtcttcatctgttgc	(ag)25	25	153	MSG0383	
MSE0250	LG04	C.M.S.	cttccccaaaccaccatcaaaata	gaaattgaagaacacgaacctgcc	(atc)3, (ca)3, (tc)19	19	340	AB623937	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSE0226	LG04	C.M.S.	tcagttccgttttgtttccttggt	tgaacctactctcatcccaaatgc	(ta)11	11	217	AB623927	
MSG0742	LG04	C.M.S.	atagtgagccatttggtgcgtttt	gtttaggttttcgcttggctatcacatc	(tc)17	17	296	AB624218	
MSG0530	LG04	C.M.S.	acggataaaatttttgtgtctggc	gtttcaaagtgaagctcccaagatgc	(tc)3, (tc)15, (tc)3	21	266	AB624114	
MSG0598	LG04	C.M.S.	atgagagtgtgagagaaggggagg	gtttcggttgagacgaatgatccaaaat	(ag)17, (ag)3	20	228	AB624149	
MSG0561	LG04	C.M.S.	accaatggttcaggcgttaaaata	gtttcaatgggtgacaacgatggtagaa	(tc)14, (ct)3, (acc)3	20	247	AB624129	
MSE0003	LG04	C.M.S.	ctctctgttccccgttctgttgat	tctgaagcagaggtgaagataaggaga	(tc)21	21	286	CV066965S	
MSG0672	LG04	C.M.S.	actccatttctcatcatcctttcc	gtttgaaaccatgaaagcacacacc	(ag)18	18	298	AB624185	
MSG0162	LG04	C.M.S.	ataatatcggctggcatcgtaatg	gtttgagaatctaagggagttgcgggtt	(tc)6, (tc)4	10	137	MSG0162	
MSG0327	LG04	C.M.S.	acgttcaaaaccgtaatttggtcc	gtttgccacagttgcaatcaagtggtag	(tc)18, (tc)3, (cgt)3, (tc)3	27	273	MSG0327	
MSG0312	LG04	C.M.S.	agtagggagtctgctgccattagt	gtttcgtctccaacaatttcactcca	(ag)25, (atg)3	28	178	MSG0312	
MSG0673	LG04	C.M.S.	agaaaccctaaccagttacggagg	gtttaccctcaaaagttcacacactcat	(ag)18	18	286	AB624186	
MSG0444	LG04	C.M.S.	attacctcagtgtcgattgccgtt	gtttaagcctcaactcagccctttca	(ag)4, (ag)8, (tg)12	24	262	AB624075	
MSG0213	LG04	C.M.S.	actgtgtcttctcgcagtttgtgt	gtttgcagcgactaccttcatctcttcc	(ct)3, (tc)9, (tc)16, (ta)3, (tg)8	39	222	MSG0213	
MSG0236	LG04	C.M.S.	ataccgatgtccccttctttctc	gtttctctgtcatcttcttcgagcc	(tc)3, (tc)4, (tc)19, (ac)6, (ct)3	35	179	MSG0236	
MSG0400	LG04	C.M.S.	atcatggatggctagattgctttg	gtttcaagtttgatcaaggggttgtatgc	(tc)15, (tc)4, (tc)3, (ct)3	25	226	MSG0400	
MSE0075	LG04	C.M.S.	tcacgttttgtagcctaggttcttca	aaatcaatgtttcgatcgctgctc	(tc)4, (tc)12	12	178	FS949483	Chapter 2
MSG0049b	LG04	C.M.S.	attggaaaatgcgtaaaaagctcg	gtttccctcgggtctttaaatatgagcc	(ag)3, (ag)20	20	152	AB623983	
MSG0380	LG04	C.M.S.	acagaccttcaccctctccatttc	gtttacctctgccttcgttcttcagc	(ag)21, (ta)4	25	277	MSG0380	
MSG0504	LG04	C.M.S.	atcccacttcttccacttcacgat	gtttcgaagagcggttgatggagaatag	(tc)16, (tc)3, (tc)3	22	201	AB624103	
MSG0642	LG04	C.M.S.	atcatgaaggacgacatgcagaag	gtttcttttgcctacgatacaatgccaag	(tc)19	19	285	AB624167	
MSE0238	LG04	SAK	ggagggttagggtttcgatcaca	agaccttcttgtgcatcttgcctc	(tc)7, (ct)3, (caa)4	7	298	AB623932	
MSG0494	LG04	SAK	agccatagagcaacagtgaaatcg	gtttccctaaaaccttgaagcaaccaa	(ag)18, (ag)4	22	254	AB624101	
MSG0240	LG04	SAK	acgcgatttcttgaactagctctc	gtttcgatcgcaaacataaacatccttg	(ct)3, (tc)26, (tc)5	34	203	MSG0240	
MSG0248	LG04	SAK	agtgggggtggagacgataaaaga	gtttgagcccgtttttcttcaaaatgac	(ggt)4, (ag)3, (ag)11, (tg)12, (ca)3	33	255	MSG0248	
MSE0015	LG04	SAK	aggccctccagaacctcacaaa	attgttccctcctacctccccata	(acc)3, (cca)3, (ta)5, (ta)5	5	268	CV014201S	
MSG0748	LG04	Kana-Ck17	atctttcctctctccgatccattg	gtttgaacaatcctccgaatcggtact	(tc)17	17	174	AB624221	
MSG0463	LG04	Kana-Ck17	agggtcgatgattgaagaagaaca	gtttcgaatcggatttaatacccgtgag	(ag)20, (at)3	23	254	AB624086	
MSG0240	LG04	Kana-Ck17	acgcgatttcttgaactagctctc	gtttcgatcgcaaacataaacatccttg	(ct)3, (tc)26, (tc)5	34	203	MSG0240	
MSG0816	LG04	Kana-Ck17	agagagacgcagaaatgattgaga	gtttccgattatttgagggaatgcag	(ag)15	15	289	AB624249	
MSG0640	LG04	Kana-Ck17	atatgtgttgaaacaccagcggac	gtttaacaactttcaacatccacggaca	(ag)3, (ag)16	19	279	AB624165	
MSG0722	LG04	Kana-Ck17	actcattcctacacagccgaaaac	gtttaacgcgtgtagagcgcaaaacta	(tc)4, (tc)13	17	103	AB624211	
MSG0810	LG04	Kana-Ck17	acacagtttcaaaccaatgacacc	gtttcaatgatggtgacagacgatga	(ag)15	15	284	AB624245	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSE0204	LG05	C.M.S.	cgatcgtcgttcttgaaatcccta	gcataacctggaacggtttggtaa	(ag)15	15	293	AB623913	
MSE0296	LG05	C.M.S.	ggggttctgcgatttctttctctg	atgattggggaaaggagaagagga	(gaa)13, (ga)3	13	272	AB623958	
MSE0299	LG05	C.M.S.	catctcctgaatcactcccagaca	atcagaatcgtcagcagcatcaac	(ac)5, (ta)3, (ag)6	6	270	AB623961	
MSE0156	LG05	C.M.S.	agaatcgctaaggcagagcaatca	atgaagcaacccagaacgactagg	(ag)5, (ag)11	11	226	AB623896	
MSE0313	LG05	C.M.S.	tgctatgccgcctaacaaaaactt	accaccaacaacaattcccactct	(ag)3, (ag)12	12	228	AB623968	
MSE0321	LG05	C.M.S.	tataactctgcggggagttttgga	tttgaaaaggaaggtgcaagcagt	(ag)13	13	292	AB623971	
MSG0541	LG05	C.M.S.	atactgcccacaaccacctaaaat	gtttacggacaagttttctcctaacccg	(tc)18, (tc)3	21	280	AB624124	
MSG0788	LG05	C.M.S.	agtcgaatatggggttaggtgtga	gtttacagaccactccctcccttca	(ag)3, (ag)13	16	300	AB624233	
MSG0446	LG05	C.M.S.	attgcccgcagtctgtttttctat	gtttgacaacgaaaagacgatctgctca	(tc)4, (tc)20	24	266	AB624076	
MSG0571	LG05	C.M.S.	agaactaggggtagaaagcgggta	gtttatgcaagcaaagggtaggttgg	(ag)20	20	195	AB624133	
MSG0702	LG05	C.M.S.	atgttggaccagtagcaccgaaac	gtttcgggttcctttctcaaactcc	(ag)18	18	283	AB624201	
MSG0578	LG05	C.M.S.	aggtgtatgggaatcgagatgagc	gtttcagtccaaattaccgacatccaca	(tc)20	20	191	AB624138	
MSG0417	LG05	SAK	acaccctcattcttctccctttt	gtttgtcggatttggactctttggtt	(tct)3, (tc)18, (tc)3	24	290	MSG0417	
MSG0435	LG05	SAK	agttaatgaccggacaaacctcca	gtttcacaacaacaacagccattcatca	(ag)3, (ca)3, (ag)18	24	281	AB624070	
MSE0293	LG05	SAK	aatgtcgatctctgcaacaccaag	gcgtcgaagaagatgaggagagag	(tc)5, (tc)4, (tc)5, (cct)3, (cct)3	5	261	AB623955	
MSG0597	LG05	Kana-Ck17	atgaatagggtcgcaggacttgaa	gtttacttagggctccgcttttcatca	(ag)17, (ag)3	20	250	AB624148	
MSG0838	LG05	Kana-Ck17	atctacggcacacaaattacgacg	gtttaaatgaaggcagtggagcaa	(ag)3, (ag)12	15	213	AB624256	
MSG0121d	LG05	Kana-Ck17	gagcaaagaagttgtgatttcgcc	attttgagatttcgaagctccacc	(ag)11	11	264	AB623989	
MSE0209	LG05	Kana-Ck17	taacagagcagaaccccttctca	gtggcagtctgtgagctgttgttt	(tc)8, (tc)4	8	299	AB623918	
MSE0167	LG06	C.M.S.	ggtctgaaaccattttcgcttctc	catcccatcactgatcatctccac	(tc)12	12	214	AB623901	
MSG0096a	LG06	C.M.S.	taaatccccgaaatccctaaatcg	tttgttccttcggagtatgatcgg	(tc)17, (ac)10, (at)3	17	234	AB623986	
MSG0423	LG06	C.M.S.	actccatgtgctgctctgtagttc	gtttgcaggaagttgagccagac	(tc)4, (tc)13, (ac)7	24	163	AB624063	
MSG0308	LG06	C.M.S.	accgtttgcaggagacatcatttt	gtttctgcctctctccctctcacta	(ag)11, (ag)8, (ag)3, (ag)3, (ag)3	28	250	MSG0308	
MSG0290	LG06	C.M.S.	agatccctgttagattgggttcaa	gtttaagatttgaaaagagggtcgaggc	(ac)3, (tc)26	29	189	MSG0290	
MSG0482	LG06	C.M.S.	attgcctcaaaaccaaggttgaag	gtttagttctgcgctggcaaagagtt	(cg)3, (ag)20	23	245	AB624097	
MSG0420	LG06	C.M.S.	acgtcaatcatgcctttgacagtt	gtttggttctgtcacggaattgaa	(tc)4, (tc)20	24	279	AB624061	
MSG0713	LG06	C.M.S.	acacagttttctcagccatttcct	gtttgtagagagcttgtgaacgcagc	(tc)3, (tc)14	17	296	AB624205	
MSG0563	LG06	C.M.S.	accatcgtagccccctaccactat	gtttcaaaccctatttaccaactgccg	(ag)17, (gat)3	20	190	AB624131	
MSG0818	LG06	C.M.S.	agcatttcagcctgtaggagtgac	gtttcaaggtaggctttagagagtaaggctc	(tc)15	15	136	AB624250	
MSE0119	LG06	C.M.S.	gcccgaagagatgttcaagtttgt	tttccatttccacctacttcccaa	(tg)3, (ag)10	10	103	AB485973	Ogino *et. al.* (2009)
MSE0121	LG06	C.M.S.	ccgctcgctaactacgactctctc	cggcaagtatgatttttccaggag	(ag)13	13	177	AB485974	Ogino *et. al.* (2009)
MSG0398	LG06	C.M.S.	atagtcagatttgtgggttcgcaa	gtttaaaacgaagattggctccagtgag	(ac)3, (tc)12, (tc)4, (ct)3, (ac)3	25	265	MSG0398	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0706	LG06	C.M.S.	attgaaaggcacatcgatcctcat	gtttgcagagatgatgagcgtaggtcaa	(tc)15, (ac)3	18	293	AB624203	
MSG0395	LG06	C.M.S.	aggtcagagatcggagatggtgac	gtttgaacaaaccctagtcaaaccc	(ag)3, (ag)6, (ag)16	25	176	MSG0395	
MSG0532	LG06	C.M.S.	agacaatctagtgcggcaatctcc	gtttccagaaatccctttttcaatctgc	(tc)4c, (tc)17	21	206	AB624116	
MSE0294	LG06	C.M.S.	ggtcctctcacctttatcttcaact	gtgcctgaagagcagacagtgaaa	(tc)18	18	249	AB623956	
MSE0047	LG06	C.M.S.	tgatcatctctaaacccctaaaaccc	tgatgttggaatggttgaaggaga	(tc)13, (tc)4, (tc)3	13	268	AB485970	Ogino *et. al.* (2009)
MSE0266	LG06	C.M.S.	aaggaaagcaggcgttcacga	aatcagtgtcggaagtgggcttag	(ag)12, (ag)3	12	207	AB623943	
MSE0154	LG06	C.M.S.	tgtgccgtagagaatggcagttta	acacatgtctgtgtagccggaaaa	(tc)13	13	323	AB623895	
MSE0035	LG06	C.M.S.	caccttcaatcttccattgacgc	cacaaaaaccccaaaataccctcc	(tc)13, (ag)4, (ag)5	13	213	AB461372	Ogino *et. al.* (2009)
MSG0233	LG06	C.M.S.	atttctcctccagatacgacggct	gtttatgggctctgacagtatggg	(agg)3, (ag)27, (aag)3, (gca)3	36	273	MSG0233	
MSG0679	LG06	C.M.S.	aggaagtccccagtaatctcagcc	gtttgcacactgagctgcatcttta	(ag)18	18	265	AB624189	
MSG0800	LG06	C.M.S.	atcttgttttgaagtgtcggtgct	gtttaacagcagcaaatacgcacaactc	(tc)16	16	181	AB624241	
MSG0472	LG06	C.M.S.	atccgaatgagagatccgaatgat	gtttgcccgaaattgataaaggtggttt	(tct)3, (tc)17, (tc)3	23	215	AB624091	
MSE0072	LG06	C.M.S.	attcagtccaccctacccccctaaa	agcatcctttggatacccttctgg	(tc)14, (tc)3	14	188	AB361049	Chapter 2
MSG0610	LG06	C.M.S.	acagaggaggaagatgatcggtaa	gtttgaagaagaagaaaactcccgccat	(tc)16, (tg)3	19	293	AB624154	
MSG0311	LG06	C.M.S.	aggtaaggactcactggtgtctcg	gtttatgtggtgattttcgaatgtcgg	(ag)10, (ag)18	28	226	MSG0311	
MSG0560	LG06	C.M.S.	accaagacacctaactcaagcgaa	gtttgcaactggttcaactcatcatcc	(at)3, (ag)17	20	178	AB624128	
MSE0138	LG06	C.M.S.	ttttccccaatttcctcaacctct	gtgtcgattgatcaacaatatgagc	(cct)7, (atc)3, (tgt)3	7	214	FS949402	
MSG0201	LG06	C.M.S.	atcatttccaaccacgcttttctc	gtttccctattctatccacattgtaccca	(tc)3, (tc)5, (tc)3, (tc)4, (tc)6, (tc)3, (tc)4, (tc)4, (tc)12	44	279	MSG0201	
MSG0670	LG06	SAK	acgaaaagcaagctacaccactcc	gtttcactcctttcaagctcttgatctc	(ag)18	18	154	AB624184	
MSE0326	LG06	SAK	atatggaaggagcaacagaaccca	aaagcctgttctgaccacaaaacc	(tc)9, (tg)4	9	241	AB623975	
MSG0121c	LG06	SAK	gagcaaagaagttgtgatttcgcc	attttgagatttcgaagctccacc	(ag)11	11	264	AB623989	
MSE0220	LG06	SAK	cacatacagcaagctcactcactca	tgagtagacccaagtagggttcgg	(ac)3, (ag)5, (ac)3, (ag)4	5	237	AB623924	
MSG0443	LG06	SAK	atggtgctagctgtttctccatgt	gtttcaagtaagcaccctcaagagctga	(tc)15, (tc)6, (ac)3	24	298	AB624074	
MSE0041	LG06	SAK	agacagtgctccatctccatctcc	tacccttggattttgcccacatac	(tc)4, (tc)4, (tc)10, (tc)3	10	283	FS950273	Chapter 2
MSE0044	LG06	SAK	cagcttagcagccatcattcaaca	actctttttacgaacaacacccgc	(ta)11, (ag)4, (ag)6	11	131	FS948949	Chapter 2
MSE0198	LG06	SAK	aacgagcctgataacaggttttgg	Tattcttcctgtgcacatcctcca	(ag)4, (ag)3, (ag)3, (ag)3	4	231	AB623911	
MSE0050	LG06	SAK	tttgagtccgtcattgttccacac	gttttgatcgagggagtcagagga	(ttc)3, (tc)14, (ca)3	14	269	FS947935	Chapter 2
MSE0011	LG06	SAK	gggggaacaacgaacagattagag	tggattagggctgaaattagggct	(ga)3, (ag)7, (ag)5, (ag)3	7	211	CV014421S	
MSG0433	LG06	SAK	agggtaggagtagttctccatcgg	gtttaaaaagatctccccactttctccg	(ag)21, (ga)3	24	152	AB624069	
MSE0053	LG06	SAK	gaaatccccaaaatttgtctcagtg	cttcacgatgagagtgtgaccgtt	(ga)3, (tc)16	16	270	AB461367	Chapter 2
MSG0619	LG06	SAK	acgcccctactgtatctggttctc	gtttacagtgaccaccgcctctacct	(gca)3, (ag)16	19	253	AB624158	
MSE0302	LG06	SAK	ttgattccatttggagagggagaa	aagaggaggagtaggaggaggcaa	(ag)3, (ag)9	9	347	AB623963	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0454	LG06	Kana-Ck17	actcattatgctttgccacctttt	gtttcaaatgatgaatggtatgccgac	(tc)3, (tc)20	23	232	AB624081	
TMS_15	LG06	Kana-Ck17	gcccccatgtatgtttctcctctga	gttgccgattgcgaagagaagataga	(tc)13, (tc)4	17	99	AB551791	
MSG0535	LG06	Kana-Ck17	agcattttggttaagtggcgttta	gtttgggaacacaacgaaggtgagctat	(ga)3, (tc)15, (tg)3	21	297	AB624119	
MSG0558	LG07	C.M.S.	acagactcatccaacctttcctct	Gtttctagtgtcttcgatttctca	(tc)3, (tc)17	20	285	AB624127	
MSE0162	LG07	C.M.S.	cctctgcttcctctgattcgtacc	Cactcagatttctttttcccacgg	(ag)16	16	208	AB623900	
MSE0170	LG07	C.M.S.	acaaccatccaccgtatatccacc	acacaaatatgcaccatcaccgtt	(tc)12, (ta)3	12	328	AB623903	
MSE0008	LG07	C.M.S.	acaagctcgggcaaccaccat	Aacaatgtctggcttcttttccca	(ta)4, (ag)14	14	129	CV013757S	
MSE0190	LG07	C.M.S.	tgtgcatcttaggacaagctcagg	gacttcttaagagcagcagccagc	(ta)4, (ag)10	10	224	AB623908	
MSE0108	LG07	C.M.S.	agtccatggtggttgatgatcctt	ttgggagtaggattcttgcagagc	(tc)6, (ta)8	8	266	FS948805	Chapter 2
MSE0239	LG07	C.M.S.	gggagcacaagaaaacaaggaaca	aagacattttatgcgccacaatcc	(ag)3, (ag)9	9	248	AB623933	
MSG0744	LG07	C.M.S.	atccatgttaaagcgactgtcacg	gtttccatctggtggtggtgatgtagt	(tc)3, (tc)14	17	283	AB624219	
MSG0133	LG07	C.M.S.	agctgaagccgtctatgatccagt	Caattgtctctggttcggttgttg	(tc)14	14	235	MSG0133	
MSE0295	LG07	C.M.S.	gaatccacttgtgttcttcatgcg	atattctccattgcaccccaaaac	(ag)13, (ag)4	13	350	AB623957	
MSG0668	LG07	C.M.S.	acctgagtgacctctcgagttcgt	gtttaaaacccagcctcccatttttc	(ag)3, (ag)15	18	216	AB624183	
MSG0403	LG07	C.M.S.	atgatcgccggtttagagatgaat	gtttaagctggctaacctacacggagc	(tc)25	25	298	MSG0403	
MSG0540	LG07	C.M.S.	ataccaagcacaatggcttctctc	gtttgacgaagtttggttccacagacct	(tc)3, (tc)14, (ac)4	21	255	AB624123	
MSE0037	LG07	C.M.S.	tatacccatttgccttgcttggtc	gaagcaaatttgggaaatcagtgc	(ca)3, (tc)12, (ag)4, (gaa)3	12	215	AB485967	Ogino *et. al.* (2009)
MSG0603	LG07	SAK	attcgcttccaaaacctccatcta	gtttcaccattgtgcacttaaccaccat	(tc)17, (cca)3	20	296	AB624150	
MSE0217	LG07	SAK	ggagatctgttctgttcggagcat	ccatcgaccatcgaatttacatca	(ag)9, (ag)3	9	344	AB623923	
MSE0241	LG07	Kana-Ck17	acatcaaaccctcttcctcctcc	Agtctttgctcccaacaaaaca	(tg)3, (ta)9	9	340	AB623934	
MSE0325b	LG07	Kana-Ck17	ggaggttcaactccagtttctcca	tttgacgtcactgcgaaatctcat	(at)3, (tc)8, (tg)3, (ta)3, (ca)3	8	324	AB623974	
MSG0643	LG07	Kana-Ck17	atccccaaagcaactcaacagaaa	gtttgtgcttctgaagtcaggatacgc	(ag)19	19	165	AB624168	
MSG0389	LG07	Kana-Ck17	agaagcaacgaagatccactgaag	gtttctatccttcgtggcaatggattt	(ag)22, (ag)3	25	292	MSG0389	
MSG0766	LG07	Kana-Ck17	accaagggcatcctatacaatgtt	gtttgccctcaatgggtttaattgagt	(ag)16	16	149	AB624228	
MSE0237	LG08	C.M.S.	ctctccttcttcacaccctccaaa	ttgttctcaaagaacctccttcgc	(acc)3, (tc)4, (tc)3, (tc)4, (tct)5	5	338	AB623931	
MSG0330	LG08	C.M.S.	agcaaatgatgatgattacgaacc	gtttcattgagttgttggcttcttcg	(tc)3, (tc)17, (tc)3, (tc)4	27	277	MSG0330	
MSG0650	LG08	C.M.S.	atgatgagaacggtaggaggtggt	gtttcgcctctttctaaatcacaaaccc	(ag)3, (ag)16	19	216	AB624173	
MSG0221	LG08	C.M.S.	agtgaatctagccgacctgcttgt	gtttgtccataaagacatgaggaggca	(agg)3, (gga)4, (ag)3, (ag)17, (tg)3, (ta)3, (ta)4	37	296	MSG0221	
MSG0237	LG08	C.M.S.	atcatgcgaagaatgaagacaacc	gtttcagcatcctatgctcctccaaaac	(ag)16, (ag)19	35	234	MSG0237	
MSG0576	LG08	C.M.S.	agcctcaagtaacctcagctcaaa	gtttccagctgtttctacaaagctccac	(ag)3, (ag)13, (ag)4	20	263	AB624136	
MSG0584	LG08	C.M.S.	ataaagaaaatggtgggttgtcgg	gtttgctgcctgtgcaacaagatta	(tc)17, (at)3	20	195	AB624141	
MSG0824a	LG08	C.M.S.	agggtggatttttaggtgtggaat	gtttggacaaggaaatgatggacaaagg	(ag)15	15	165	AB624252	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0127	LG08	C.M.S.	atcgacagatctcgaaacagacca	gtttacacaacactccctcacttcctcc	(ag)11	11	184	MSG0127	
MSG0720	LG08	C.M.S.	actaaggtccctcaacctcataccc	gtttggttggggaatctagggtttc	(ct)3, (tc)14	17	206	AB624210	
MSG0359	LG08	C.M.S.	actctcatttgcaaccacactttg	gtttcccacattgtacccacattcacat	(tc)3, (tc)5, (tc)7, (tc)7, (tc)4	26	228	MSG0359	
MSG0316	LG08	C.M.S.	atgaccacccctcttgaaattagg	gtttagctgagatttgagctcaggaacg	(tc)21, (ac)3, (tc)4	28	197	MSG0316	
MSE0135	LG08	C.M.S.	aagatccaggtcagaccagcc	aatgttgagagcgaatgagggaag	(tc)3, (ag)10	10	204	FS950401	
MSE0100	LG08	C.M.S.	ttctttccgtgtacatacccccc	gaattgttggaggccgtagaattg	(tc)15	15	253	AB461370	Ogino *et. al.* (2009)
MSE0271	LG08	C.M.S.	ggctcttcttagatccatacaaacca	acctcatctgagtcctcccatgtc	(tc)13	13	340	AB623946	
MSG0661	LG08	C.M.S.	acaacccgaatttaaccggtattt	gtttacatccaccacagaacaatggcta	(tc)15, (tc)3	18	165	AB624179	
MSG0413	LG08	C.M.S.	attgtcgatccaaccacaatcg	gtttcttggtgtgttctctgaggctg	(tc)25	25	290	MSG0413	
MSE0304	LG08	C.M.S.	ggaaaccctagctcttcagcacaa	ccaaaccccaacaaacataacagc	(tc)7, (tc)5	7	329	AB623964	
MSG0700	LG08	C.M.S.	atgggtattgcaaggaaaagagga	gtttagagacagttgaccgggcttc	(ag)14, (ag)4	18	198	AB624200	
MSG0821	LG08	C.M.S.	agcgaatttcagtgcgtctatctc	gtttcgtcatcgtcttcaatttggtgt	(tc)15	15	239	AB624251	
MSE0215	LG08	C.M.S.	tagagggctttgacacagcaactg	ggtggatatcatccaaccttccac	(tc)10, (tg)5	10	227	AB623922	
MSG0824b	LG08	C.M.S.	agggtggatttttaggtgtggaat	gtttggacaaggaaatgatggacaaagg	(ag)15	15	165	AB624252	
MSG0808	LG08	C.M.S.	attctgggtttattttgagggcgt	gtttgatagtagatcggttgcgacgaa	(tc)12, (tc)4	16	191	AB624244	
MSG0779	LG08	C.M.S.	agaggttcgtgcaaggtcaacatc	gtttggcttgcttatgtatccctg	(tc)16	16	281	AB624230	
MSE0159	LG08	C.M.S.	aaatcttgatcgcaatcttgctcc	caaattgatttctcacatagggcca	(tc)20, (tg)5	20	298	AB623898	
MSE0291	LG08	C.M.S.	aatcaaataacacttgcacccgc	aaaaagagaaagtcacgtccacgg	(ca)3, (tc)9	9	231	AB623953	
MSG0388	LG08	C.M.S.	actttgtcttcttcgtcgttgtga	gtttaacgcctccaaatgtggtagacac	(ag)14, (ag)4, (ag)3, (ag)4	25	183	MSG0388	
MSG0587	LG08	C.M.S.	ataatcgagttgaggcaaggcaag	gtttccctatggtttgtgactgatgtgc	(ag)20	20	242	AB624142	
MSE0089	LG08	C.M.S.	gccgtcttcttcttcttctctccc	ataatcacctggaaaacagcggaa	(ag)5, (ag)7, (tg)3	7	294	FS949251	Chapter 2
MSG0353	LG08	C.M.S.	acgattaagggagagaagagaggg	gtttccaatccacaccagactctt	(ag)4, (ag)4, (ag)15, (ag)3	26	221	MSG0353	
MSE0276	LG08	SAK	gtcaagatcgggtctagggttcct	ccactcagcgacaattctctcctt	(tc)14, (tc)3	14	328	AB623948	
MSG0659	LG08	SAK	attttcacgggcagctctctctct	gtttagttggggaaaatcgaaaatggg	(tc)16, (ttc)3	19	284	AB624177	
MSE0017	LG08	SAK	ttctccaccctctgctatctctcg	caataccatgtgcccaaaaacaga	(ac)3, (ta)3, (ag)4, (ag)6	6	216	CV014903S	
MSG0646	LG08	SAK	atcggaaaccaaacaggacgatta	gtttacaatcgacgatgaagctcgtg	(ag)16, (agt)3	19	250	AB624170	
MSG0525	LG08	SAK	acaccacacccatctctatcccat	gtttccaattccagtcaccctcctttta	(tc)18, (agg)3	21	236	AB624111	
MSE0051	LG08	SAK	ccttgcaagtggacaaaactgttc	agctgcagagacgtaggcatttct	(tc)15, (ta)4	15	189	AB461365	Chapter 2
MSE0031	LG08	SAK	cagatctgaaggtgaagggtttcg	tttcttatgagacttggccttcgc	(tc)8, (tct)9, (ac)3, (ca)3	9	269	FS951496	Chapter 2
MSG0142	LG08	SAK	atggagaagggtgagatttgagga	gtttaccgatctaccaccatcacctcat	(ag)10	10	146	MSG0142	
TMS-LA3	LG08	SAK	cctctcctcccctcccaaaa	ttcctccgtcgttgtctctcc	(ga)5, (ca)9	14	151	AB212680	
MSE0230b	LG08	SAK	agtctagttgggcaacggtgagac	aaaagtcattcccctttcccctt	(tc)4, (tc)4, (tc)5	5	348	AB623929	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSE0028	LG08	SAK	ccaacccaaaatttcacatgatagc	ttcaacaacatcgctagaggtcca	(tc)21, (ag)3	21	177	FS951387	Chapter 2
MSE0288	LG08	Kana-Ck17	gcagagatggaggggactctgtaa	tcgccagaactcaacagaacactc	(tc)10, (tg)3	10	266	AB623952	
MSG0777	LG08	Kana-Ck17	agacctagaaattcttcacttcccc	gtttagggttcgaaagggagactggt	(ta)3, (ag)13	16	258	AB624229	
MSG0203	LG08	Kana-Ck17	attccccttcttaacagccacaga	gtttaggggtagacgagtgctcaacatc	(gat)3, (tg)4, (tg)19, (tc)14, (ac)4	44	255	MSG0203	
MSG0525	LG08	Kana-Ck17	acaccacacccatctctatcccat	gtttccaattccagtcaccctccttta	(tc)18, (agg)3	21	236	AB624111	
MSG0629	LG08	Kana-Ck17	agccgatacagtcggatacatctc	gtttcacccatttcaatctcagactggt	(ag)19	19	271	AB624162	
MSG0412	LG08	Kana-Ck17	attggttactagccgttggtagcg	gtttacaacctctggcctctctttaccc	(ag)25	25	269	MSG0412	
MSG0142	LG08	Kana-Ck17	atggagaagggtgagatttgagga	gtttaccgatctaccaccatcacctcat	(ag)10	10	146	MSG0142	
TMS-LA3	LG08	Kana-Ck17	cctctcctcccctcccaaaa	ttcctccgtcgttgtctctcc	(ga)5, (ca)9	14	151	AB212680	
MSG0730	LG08	Kana-Ck17	agggagatggaacatttctcgaat	gtttctttcctctcccttttctcgga	(gt)3, (ag)3, (ag)8, (ag)3	17	231	AB624214	
MSG0731	LG09	C.M.S.	aggttgaaatgccagaaacgagag	gtttcgacatcgcctgttcttcacttta	(ag)9, (ag)8	17	199	AB624215	
MSE0181	LG09	C.M.S.	cggtttctctgaagacagatttcacat	ttctcacaccgcgaaacctaatct	(tc)8, (tc)4, (tc)4	8	238	AB623907	
MSG0426	LG09	C.M.S.	agatcgcttgagactggattgatt	gtttgaagaagatgaggagcccatcaaa	(tg)3, (ag)21	24	280	AB624065	
MSE0260	LG09	C.M.S.	tgaattaatgggtgagatgcctga	cgatgcatacaaagattgagccac	(ac)4, (ag)15	15	246	AB623941	
MSG0300	LG09	C.M.S.	atgatttggagatttggtttgtagg	gtttgaagaaaaaggtaagggagaaaggga	(tc)14, (tc)15	29	284	MSG0300	
MSE0113	LG09	C.M.S.	taccttctgcaactccagcaatcc	tgagattgaccatctttcatcgga	(tc)14	14	223	AB485972	Ogino *et. al.* (2009)
MSG0297	LG09	C.M.S.	atcgagtggctcgactttgatttt	gtttaaacccgaatctggactaccacaa	(ca)3, (tc)20, (tc)6	29	173	MSG0297	
MSE0342	LG09	C.M.S.	gctggtcagtttagaaatgaagatg	tgtgagcaaaccctaccctaccat	(tg)15	15	292	AB623980	
MSG0760	LG09	C.M.S.	acaaatacgtagcagacatgaagca	gtttaagtttgggggacaaacattgaga	(ag)16	16	185	AB624225	
MSG0511	LG09	C.M.S.	atgatgagctcccttcaaatgtca	gtttagactaatttgagtcgacccgacc	(tc)22	22	296	AB624105	
MSG0574	LG09	C.M.S.	agccagctgtttgcacagatttat	gtttagggctcaaagaaggtcccaat	(tc)20	20	282	AB624135	
MSG0452	LG09	C.M.S.	acctctgcctttcgatcactgtaa	gtttgttcttgttcccaacctccac	(tct)6, (tc)14, (at)3	23	238	AB624080	
MSG0657	LG09	C.M.S.	attcaggtgacatggtccatcctt	gtttagcatgatgacctacaatgtgtg	(tc)19	19	180	AB624176	
MSE0101	LG09	C.M.S.	atgctctcctctgtttacgccttg	acgaattctgaccacgaaacccta	(gc)3, (aac)3, (cca)6, (ccg)3	6	250	FS950763	Chapter 2
MSE0306	LG09	C.M.S.	caccaacgaacggttcagataaca	cttatgttggcgatcgggagatac	(tc)3, (tc)5, (tc)8	8	330	AB623966	
MSE0022	LG09	C.M.S.	cctgcagtgtagaaaagccccaat	cgatctgggagcttcttggagata	(tc)19, (ta)9, (ac)5	19	192	AB361048	Ogino *et. al.* (2009)
MSG0577	LG09	C.M.S.	agggattggagttctagggttttg	gtttacagctacggtcggaagagaacc	(ag)11, (ggc)3, (ag)3, (ga)3	20	174	AB624137	
MSG0649	LG09	C.M.S.	atgagaagcgtgtttgcactagga	gtttaagagagacgaaccaatcgggttt	(tc)19	19	300	AB624172	
MSE0205	LG09	C.M.S.	cggccttcccgtcttatcttcata	catttccgaaggagtggttgaatc	(ac)3, (tc)11	11	286	AB623914	
MSE0171	LG09	C.M.S.	attcatccatatccatgccgacac	gaaacgtgcgaaaaagaaatggtc	(tc)16, (ctc)3	16	213	AB623904	
MSE0012	LG09	C.M.S.	gggagaaccaacccagtctatactcc	gcttggaaagttcggagacgatta	(tg)3, (ag)14	14	291	CV013680S	
MSG0421	LG09	C.M.S.	acgtcccagtggatcaaaagaata	gtttgcttccaccaatgatcgtaaaga	(ct)3, (tc)15, (ct)3, (tc)3	24	288	AB624062	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSE0143	LG09	C.M.S.	gccttttgtcagaaacggtgact	cagcaatctttggttttgttgtgc	(ag)8, (tc)4	8	176	FS951913	
MSG0719	LG09	SAK	acgaagaaaaggtcagccctaaaa	gtttggacacttcctcaattagggtttgc	(ag)14, (gga)3	17	300	AB624209	
MSG0467	LG09	SAK	ataaaacgtttggttcttcactcac	gtttaagagaaggcgaagagggatgact	(tc)20, (cca)3	23	172	AB624087	
MSG0798	LG09	SAK	atcgtagcacacaaacacgatgaa	gtttgggaaaacctaatgtgcagc	(tc)3, (tc)10, (tc)3	16	298	AB624239	
MSE0298	LG09	SAK	cctccatttctgcttcaatcacct	gtctcatcaacacacaatccccag	(tc)6, (ct)3, (tc)4	6	239	AB623960	
MSG0439	LG09	SAK	atcaagggaccctgcactatgaaa	gtttcattcctccattcccttctcc	(ag)17, (ag)4, (ag)3	24	153	AB624073	
MSE0032	LG09	SAK	tttttccctcaaaccttcttcagc	ttaatcccctggtgctcattctgt	(ag)4, (ac)13, (tc)3, (tc)3	13	221	FS949441	Chapter 2
TMS-LA4	LG09	SAK	gatctacagccacggttctgag	tcaacacattctacaaattcatcatc	(at)7, (ac)13	20	175	AB551792	
MSG0490	LG09	SAK	agaagcggcagcgtatagatcttg	gtttcgtctcacgggtgtggatctt	(ag)3, (ag)3, (ag)13, (tg)3	22	139	AB624100	
MSE0055	LG09	SAK	atcacgaaccacctcctcctcat	tgcagatatacataactccaaccct	(cca)3, (acc)3, (cca)3, (ta)6, (ta)4	6	254	FS949373	Chapter 2
MSG0641	LG09	Kana-Ck17	atcagtgctcttagatgttgcgga	gtttggaaaacttggaccgttcatct	(ag)19	19	241	AB624166	
MSE0292	LG09	Kana-Ck17	ggggttcttttagctatccccc	actcagaccattgcacgattcaaa	(tc)12, (ta)3, (ct)3, (tc)4	12	345	AB623954	
MSG0480	LG09	Kana-Ck17	attaccagatgggaggggaaaaga	gtttgagctcagtcagtctcaatcgcac	(ag)5, (ag)18	23	240	AB624095	
MSG0469	LG09	Kana-Ck17	ataccctgaggctgtgagaagttg	gtttaagttctgccaaacaggcaca	(ag)23	23	138	AB624088	
MSG0404	LG09	Kana-Ck17	atgcttggtagatggtggtagtgg	gtttgcatgtagcaccaccattttgt	(ag)25	25	235	MSG0404	
MSG0593	LG09	Kana-Ck17	atatctttgtccaccggcaagttc	gtttgaaaattgtaagccactagggcaa	(tc)20	20	286	AB624145	
MSG0282	LG09	Kana-Ck17	atcacggaccagggttatcacagt	gtttcatacaaaatgggtttcgggtc	(tg)3, (ag)14, (ag)5, (ag)4, (ag)4	30	272	MSG0282	
MSE0212	LG10	C.M.S.	ccacacacacaaacacgactcatc	gcttctttctgcagatcgtccttc	(tc)13	13	227	AB623920	
MSE0327	LG10	C.M.S.	tcttcttctcacttgtcttcaaaccc	aggcaaattccttgagatcacgac	(tc)7, (tc)4, (ac)3	7	279	AB623976	
MSG0206	LG10	C.M.S.	atgagaaggtcatccaccatgaaa	gtttaaaaatccacatcccctcgttcc	(ag)8, (ag)10, (ag)3, (ag)19, (tat)3	43	260	MSG0206	
MSG0451	LG10	C.M.S.	accctgctgtttgtttggtagaga	gtttcacctatgctctgccacactctgt	(ag)5, (ag)18	23	245	AB624079	
MSE0202	LG10	C.M.S.	ttcttcccattcccacaaaagagg	attgaaagatggagcgaaaccaga	(ag)9, (ga)3, (ag)3	9	276	AB623912	
MSG0626	LG10	C.M.S.	agaggggtagagatgaagcagcct	gtttgagaggcctccatgttagtgctgt	(ag)19	19	150	AB624160	
MSG0538a	LG10	C.M.S.	agtccacaaggacatgttcagaca	gtttaacagagacgatgaaatgacgacg	(tc)5, (tc)16	21	203	AB624121	
MSE0149	LG10	C.M.S.	caaaccagaattaccagctcatatcc	tgggtatttggagacagcaacaga	(tc)12	12	215	AB485975	Ogino *et. al.* (2009)
MSG0594	LG10	C.M.S.	atctctctcatcgtccgatcttgg	gtttcaagcaaatcgcaacatttataccg	(tc)6, (tc)14	20	296	AB624146	
MSG0604	LG10	C.M.S.	attggacttattctttcccccgac	gtttgaagatctaggtgagaaaagttgag	(tc)20	20	178	AB624151	
MSG0688	LG10	C.M.S.	ataggttagtggcaattgaggggt	gtttgctttccaccatacctgtttt	(tc)18	18	245	AB624194	
MSG0623	LG10	C.M.S.	acttcttgcgctcgagtggactat	gtttgaggctaaggtgaggaatgagggt	(tc)19	19	189	AB624159	
MSG0681	LG10	C.M.S.	agggtttgcgtcttcaaagagaga	gtttgtaacacttgccacgtttcg	(ag)18	18	235	AB624190	
MSG0652	LG10	C.M.S.	atggaacgagggaagaatgttgat	gtttactcccacttggttcgaaaac	(ag)13, (cg)3, (ag)3	19	228	AB624174	
MSG0019	LG10	C.M.S.	atgcaacaatggagagcaagagtg	gtttaagcactacaaagcacagcacctg	(ag)7	7	287	AB623981	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0686	LG10	C.M.S.	atagagatcgtgcagcgagaaagt	gtttcactcgggtgaaaaatccaatc	(ag)11, (ag)3, (ag)4	18	278	AB624193	
MSG0476	LG10	SAK	atggagggtcgcctatgatggagt	gtttggagggagacttttgatttttggc	(ga)3, (ag)14, (ag)6	23	290	AB624093	
MSG0684	LG10	SAK	atacgtgaaattgtggaagaacca	gtttgagaggaaaccgtgagaaagggtt	(ag)3, (tc)15	18	259	AB624191	
MSE0042	LG10	SAK	agactcgctcgctctcaacaaac	aagcggagataatggtggttagca	(tc)16, (tc)5	16	114	FS950639	Chapter 2
MSG0140a	LG10	SAK	atgaccctctctttgttgccaag	tgccttaagagctaaaatgccagc	(tc)3, (tc)5, (tc)3, (tc)19	19	265	MSG0140	
MSG0536	LG10	SAK	agctctccagctgatcactaaacg	gtttcaccttcaactgtcgagtcca	(ac)3, (ag)15, (ag)3	21	174	AB624120	
MSG0460b	LG10	Kana-Ck17	agatgaaggggacaacccaatttt	gtttcaatggttgagattgtttgctcac	(ag)20, (gaa)3	23	151	AB624084	
MSG0443	LG10	Kana-Ck17	atggtgctagctgtttctccatgt	gtttcaagtaagcaccctcaagagctga	(tc)15, (tc)6, (ac)3	24	298	AB624074	
MSE0287	LG10	Kana-Ck17	atttttccccaacacgcacact	ttcttgaagagccctcgtcttctc	(tc)6, (tc)3, (tc)4	6	209	AB623951	
MSG0505	LG11	C.M.S.	atcctactgtgcgtgtcaccttgt	gtttgcactccatcaatcagcaagagaa	(tc)17, (tc)5	17	215	AB624104	
MSG0648	LG11	C.M.S.	atgaagctatggtgagatggaggc	gtttcgatcagagagatcgtcaacgcta	(ag)16, (ac)3	19	188	AB624171	
MSG0714	LG11	C.M.S.	acagagccgaagacgttttacaat	gtttgctgagtattgcttgccaaatg	(tc)17	17	213	AB624206	
MSG0204	LG11	C.M.S.	attttggaaaacgaatgaagccct	gtttcgaaaagataaagggatgctgtgg	(ac)23, (tc)15, (tc)3, (tc)3	44	198	MSG0204	
MSG0617	LG11	C.M.S.	accgagtccagaatacaacaaagg	gtttcaaacctcatttctcgctccactt	(ag)19	19	297	AB624157	
MSG0457	LG11	C.M.S.	agacaaggcagtgcaaaagtcgat	gtttctctctcaaacgtgagttgtggg	(ag)5, (ag)18	23	283	AB624082	
MSG0083b	LG11	C.M.S.	gaggaaagagattatgcggtgtgg	gtgagccttcaaaagacagcaacg	(gga)3, (tg)11, (ag)24, (tca)3	24	195	AB623985	
MSE0334	LG11	C.M.S.	cagctcctccagttgctgtgtaaa	ctcctgatgctgaatccattaccc	(ag)3, (ag)10	10	270	AB623978	
MSG0665	LG11	C.M.S.	accaaagcctatctacacgcacaa	gtttaggcggagaaaaggaggtatatgg	(tc)3, (tc)15	18	282	AB624181	
MSE0107	LG11	C.M.S.	tctctctactcctgcgcaatctca	tcaaagatgttgctctcgtcaacc	(tc)8, (ca)3, (cct)3	8	164	AB485971	Ogino *et. al.* (2009)
MSG0734	LG11	C.M.S.	agtctgcagatgagaaatcaagcc	gtttgagactgtggaacggagaagtg	(tc)17	17	229	AB624216	
MSG0529	LG11	C.M.S.	accaggggaggttaaggaagtctc	gtttctcttccaacgcgcctggtta	(ag)21	21	220	AB624113	
MSG0477	LG11	C.M.S.	atggtaataagcgcacaatctggg	gtttgccaagctcatctcactttga	(tc)20, (tc)3	23	282	AB624094	
MSG0450	LG11	C.M.S.	accaggagttgaatatgcccaagt	gtttgcccaacgtctctagcattctca	(ag)3, (tc)17, (ag)3	23	265	AB624078	
MSG0753	LG11	C.M.S.	attccccttccttgtatttgcct	gtttgcgactgaaggagagagaaacgtc	(tc)17	17	157	AB624222	
MSG0838	LG11	SAK	atctacggcacacaaattacgacg	gtttaaatgaaggcagtggagcaa	(ag)3, (ag)12	15	213	AB624256	
MSG0501	LG11	SAK	atatggacacattcaggacacgtt	gtttccaccaacaactgactagcgaaga	(ag)3, (ag)19	22	227	AB624102	
MSG0716	LG11	SAK	acatgtaaggggagggttacctga	gtttgggtggatattcttttccgatcc	(ag)17	17	199	AB624207	
MSE0314	LG11	SAK	cccacactcacacacacatctcag	gggggtacaccttcctctgtttct	(tct)11, (ttc)3	11	253	AB623969	
MSG0352	LG11	Kana-Ck17	acccctttgtcaactctttgggaat	gtttaggagaagaggaggagaaggggag	(ag)26	26	233	MSG0352	
MSE0208	LG11	Kana-Ck17	ctttggaagtccaattgttcccag	ggtgactttgtggtggatctttcc	(ag)3, (ag)5, (gaa)11	11	257	AB623917	
MSG0432	LG11	Kana-Ck17	agggtaaaatgggaattggagaaa	gtttacatatcaccatcatgcaccacc	(ag)14, (tg)3, (ag)4, (atg)3	24	224	AB624068	
MSG0562	LG11	Kana-Ck17	accacacagaagaccatccctact	gtttcactttaggtgccagttccaggtt	(ag)20	20	217	AB624130	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0785	LG11	Kana-Ck17	agggaaacagagaaagggagaaag	gtttacagccgcacatcacaggttc	(ag)3, (ac)3, (ga)3, (ag)7	16	134	AB624232	
MSE0308	LG11	Kana-Ck17	tcatctccgggaagtaccatttgt	cttgctctccactttcaggacaca	(tc)12	12	330	AB623967	
MSG0662	LG11	Kana-Ck17	acaatgatggtgtgacgacggtag	gtttggtatggggataattacagtccca	(ggt)3, (ag)15	18	277	AB624180	
MSE0335	LG12	C.M.S.	tccatcctcctaaccaaattctga	actcaaaacatcgtaggcttcgga	(ag)4, (ag)12	12	253	AB623979	
MSE0161	LG12	C.M.S.	ggggacggctaagaaacaaaccta	aagtgttcatgacattgacgcagg	(tc)11	11	212	AB623899	
MSG0674	LG12	C.M.S.	agacagaaaattccacttgtgggg	gtttacatgtagcccctgtcaccctct	(tc)6, (tc)3, (tc)3, (tc)3, (tc)3	18	295	AB624187	
MSG0325	LG12	C.M.S.	acgctctgtatcggtgaaggctac	gtttgcctcttaagctattccaaaaccc	(ag)16, (ag)4, (tg)3, (ag)4	27	275	MSG0325	
MSG0832	LG12	C.M.S.	atacttctcgaagaggatcatggg	gtttaacggctctcttcccatctcac	(ag)12, (gtg)3	15	216	AB624254	
MSE0030	LG12	C.M.S.	ttccaaaaccctagtttcactcca	acgctctgtatcggtgaaggctac	(tc)4, (ca)3, (tc)5, (tc)11	11	251	AB461369	Ogino *et. al.* (2009)
MSG0534	LG12	C.M.S.	agagaccactcacccataaccatt	gtttgtcgtgttcttcttcggtcc	(ca)3, (cg)3, (tc)15	21	254	AB624118	
MSG0126	LG12	C.M.S.	atgtaaagggacaactgggattgc	gtttccacggagttcctttcttcctc	(tc)14	14	231	MSG0126	
MSE0019	LG12	C.M.S.	gccatcaaactactctcatcccca	tcccacttctatgtcctttgccat	(ac)23, (tc)12, (ac)3, (ac)4	23	143	FS951295	Chapter 2
MSG0260	LG12	C.M.S.	acttgtcctctagtctaggcttgct	gtttgatgggtttgatgaaatgggt	(tc)19, (tc)5, (tc)5, (tc)3	32	273	MSG0260	
MSE0083	LG12	C.M.S.	ctctgatcgaaatcaccaatgcct	gaagaagagagagaagtccgtgcg	(tct)4, (ct)3, (tc)3, (tct)6	6	241	FS951085	Chapter 2
MSG0796	LG12	C.M.S.	atcgaaacttgggaggtgtttttg	gtttaccctcctttcctcttcctcctct	(ggt)7, (gga)3, (tg)3, (ag)3	16	285	AB624238	
MSG0023	LG12	C.M.S.	attgaagacagggtggtggagatt	gtttactacaagtgctcctttgcccc	(ag)14	14	290	AB623982	
MSG0488	LG12	C.M.S.	agaaacctaaaccaaaccccctct	gtttgcaacattatcattgaggctgg	(tc)22	22	224	AB624099	
MSG0627	LG12	C.M.S.	agatcgagatagatgtttccgtgg	gtttacctgtcacttctctccctcgc	(ag)16, (ag)3	19	286	AB624161	
MSG0794	LG12	C.M.S.	atatgagagaagatgagccaagca	gtttgataccaagttgtaacgccctca	(ag)16	16	224	AB624236	
MSE0027	LG12	SAK	ctccctcttacttcaatcgacccc	ttgaagggcggtacttgtatggat	(cct)3, (tc)19, (tc)3	19	133	FS952009	Chapter 2
MSE0228	LG12	SAK	ggggctaagcgaatggtagaactt	aatcacccaatccaaacaaacacc	(ag)3, (ag)7, (gaa)4	7	287	AB623928	
MSG0303	LG12	SAK	atgtgtgtgtcttgttcatccctg	gtttcgctaacatgaagtgaagtggtgc	(at)3, (tc)26	29	273	MSG0303	
MSG0633	LG12	SAK	agggattattcgacattcatccaa	gtttggatttggtatgcctccccagtat	(ag)19	19	251	AB624163	
MSG0483	LG12	SAK	atttactgagacagtcgctggggt	gtttcaacatgggcacatcttccata	(ag)20, (ag)3	23	264	AB624098	
MSE0052	LG12	SAK	tgttacggtacttcttgttagctggc	tggttggtgtgtctggtgattctt	(ac)3, (tc)4, (tc)9, (tc)3	9	271	AB505872	Chapter 2
MSG0781	LG12	SAK	agataaactcgaggggaacaagcc	gtttggcagactcagggtattttgga	(ag)3, (ag)13	16	290	AB624231	
MSG0580	LG12	Kana-Ck17	agtgttgtttgagacttgagggga	gtttcgcttagagaaggtgacttgcc	(ag)14, (ag)6	20	222	AB624139	
MSG0677	LG12	Kana-Ck17	agccacctaactaccacacaactg	gtttggagtggattgatcgacgataagg	(at)3, (ag)15	18	274	AB624188	
MSG0303	LG12	Kana-Ck17	atgtgtgtgtcttgttcatccctg	gtttcgctaacatgaagtgaagtggtgc	(at)3, (tc)26	29	273	MSG0303	
MSG0690	LG12	Kana-Ck17	atcaactgtttcatccacgtgtca	gtttgagagcttggaagtgcgttaatgg	(tc)18	18	296	AB624196	
MSE0300	LG12	Kana-Ck17	cacttcgtcccctatttttgttgc	gggtttcgatgaactgattggaag	(tc)4, (tc)4, (tc)4, (ta)5	5	224	AB623962	
MSG0685	LG13	C.M.S.	atagaatgctacggtggcaaagag	gtttgcgtgacatctccttcgttttctt	(ag)3, (ag)15	18	281	AB624192	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0102	LG13	C.M.S.	attcatccactcctcagtgtctcg	gtttctttgtctcgcttgtgctcct	(ag)15	15	288	AB623988	
MSG0708	LG13	C.M.S.	attgatggtggtggtgaatctgag	gtttccatgtgttttaatcccagca	(ag)18	18	274	AB624204	
MSE0169	LG13	C.M.S.	gagaccagacctcgaatcaaagga	gacggcagcaagaacaactactcc	(ta)4, (gat)8	8	262	AB623902	
MSG0272	LG13	C.M.S.	aggccattgttgaagaacttgtct	gtttcgggttaaaagtcaacccatctg	(tc)4, (ag)18, (ga)3, (ag)3, (ga)3	31	225	MSG0272	
MSG0795	LG13	C.M.S.	atcaagcgctttttcagtctctcc	gtttggttagggcaatcaacattcc	(tc)16	16	148	AB624237	
MSG0799	LG13	C.M.S.	atctgacgagcattcgagaagaaa	gtttaaatgccccctattacaaaatggc	(ct)3, (tc)13	16	223	AB624240	
MSE0070	LG13	C.M.S.	cacacctctcatcacccttctca	gcattcatggttgcaagagagaga	(tcc)3, (tc)8, (tc)6	8	132	AB505870	Chapter 2
MSG0140b	LG13	C.M.S.	atgaccctctcttttgttgccaag	tgccttaagagctaaaatgccagc	(tc)3, (tc)5, (tc)3, (tc)19	19	265	MSG0140	
MSG0271	LG13	C.M.S.	aggatgttgggtaaagctaaggga	gtttctgttcgggacgacactttta	(ag)4, (ag)19, (ag)8	31	182	MSG0271	
MSG0470	LG13	C.M.S.	atagggttcgaaaatggcagg	gtttgaggtggcaagtttgtgactgt	(ga)3, (ag)20	23	172	AB624089	
MSE0191	LG13	C.M.S.	ttatatcctccctcaggctcctcc	cagcaatgaaacgaaacccagatt	(tc)5, (tc)5, (tc)14, (ag)3	14	290	AB623909	
MSG0527	LG13	C.M.S.	acatctctcccaagagctgaata	gtttacatgatgcaataacaaaagcggg	(tc)21	21	277	AB624112	
MSG0765	LG13	C.M.S.	acatccaccacaagtcagagaagc	gtttggtggaaacaatgaacaaccatc	(ct)3, (tc)13	16	109	AB624227	
MSG0385	LG13	C.M.S.	acctctgcattgttgtacatggct	gtttaggaaaggcaaggaaggagactgt	(ag)12, (ag)3, (ag)3, (ag)3, (ag)4	25	201	MSG0385	
MSE0324	LG13	SAK	gaacccaattccggcaaatacata	aaactccactggctctgaatgtcc	(ag)3, (ag)10	10	261	AB623973	
MSE0328	LG13	SAK	ccgaactcacaaacctctactttca	caacatcctctgcttggaaaaacc	(tc)12, (ta)3	12	242	AB623977	
MSG0315	LG13	SAK	atccattgctctggctttacattg	gtttgctatcgcagtccctcaacttc	(ag)28	28	207	MSG0315	
MSG0792	LG13	SAK	atacgatctgagctctccaacaca	gtttgagggaatgaggtgggtgagaata	(tc)8, (tc)4, (tc)4	16	121	AB624234	
MSE0024	LG13	SAK	gaaaccatcattccaaagccacaa	ccgatttcgacatctcttcgtctt	(ag)3, (ag)11, (gaa)3, (ag)4, (ag)3, (ag)3	11	262	FS951122	Chapter 2
MSG0595	LG13	SAK	atctgaggggtgctgatttgtctt	gtttaccagggtaaggaggatttaccga	(tc)15, (tc)5	20	265	AB624147	
MSG0831	LG13	Kana-Ck17	ataccatttgacccgaccctaaac	gtttaagttggggattttaatccacggt	(tc)15	15	272	AB624253	
MSE0140	LG14	C.M.S.	ccaaccaaatggactcaactctga	ggtttcttgggaaatggaggtagg	(cac)3, (cca)6, (ct)3	6	141	FS950689	
MSG0393	LG14	C.M.S.	aggaaagattgaggcgttacaatg	gtttgaagagagcggtggcataacagat	(tc)22, (tc)3	25	259	MSG0393	
MSG0637	LG14	C.M.S.	atacattatcgcccacgttcaacc	gtttaggtatcaggcaaccctgttttca	(ag)19	19	215	AB624164	
MSG0759	LG14	C.M.S.	atttcttcttcctccgacaactgc	gtttcatgcttcgcactttccttttc	(tc)17	17	266	AB624224	
MSG0660	LG14	C.M.S.	aagagagaagttgggtgcgacaaa	gtttgcagatctggacttcttctctggg	(ta)3, (ag)15	18	179	AB624178	
MSE0259	LG14	C.M.S.	tcctcaagttcagcaaggtgtgag	gtgttgttgaaagccaggaggtct	(ag)11	11	279	AB623940	
MSG0234	LG14	C.M.S.	actttctggtttacgcaattggac	gtttcccacctctttctgtcactgt	(ag)5, (ag)5, (ag)3, (ag)4, (ag)18	35	288	MSG0234	
MSG0763	LG14	C.M.S.	acaatgctatgaagagggagggag	gtttcagttagtagcatttccttggcgg	(ag)16	16	183	AB624226	
MSG0523	LG14	C.M.S.	aacagagagtgatcggtgcagaat	gtttgatctgtctccctcacctcatca	(ga)3, (ag)18	21	266	AB624110	
MSG0436	LG14	C.M.S.	ataaacaaaaaccctaaggagccc	gtttaaagatagtgccgtcgatggaaaa	(tc)20, (tc)4	24	107	AB624071	
MSG0693	LG14	C.M.S.	atcccaatttctcctcaatcctcc	gtttgagaggagaagtgctagaccgc	(tc)4, (tc)14	18	295	AB624197	

Table 3.1. continued.

Marker name	Linkage group	Mappable material	Forward primer sequence	Reverse primer sequence	Motif	Repeat (total)	Expected Amplicon size (bp)	Accsession number	Reference
MSG0401	LG14	C.M.S.	atctccttcctccaaaaccctagc	gtttaagaaccagagaggcgatttctga	(ag)8, (ag)11, (ga)3, (ag)3	25	274	MSG0401	
MSG0703	LG14	C.M.S.	atgttgttcacgagttgaggctgt	gtttgaaccctaaccctaacccttc	(ag)18	18	159	AB624202	
MSG0263	LG14	C.M.S.	aggtttttgatagggaggatggaa	gtttctctttctctccctcatcgtc	(ag)3, (ag)5, (ag)24	32	239	MSG0263	
MSG0607	LG14	C.M.S.	aatcatctctgctctcggaggatt	gtttagactgccttttcagttcacaccc	(tc)16, (agg)3	19	222	AB624152	
MSE0062	LG14	C.M.S.	tcagagaacattgaacatcaagcc	ctaggttctttccccaatcgacct	(ag)18	18	134	AB461364	Chapter 2
MSG0570	LG14	C.M.S.	agaaataacacggtatccggtcaa	gtttgtgattgcaacaccaaactcctt	(ag)17, (ag)3	20	177	AB624132	
MSG0407	LG14	C.M.S.	atgtgaaaaagtaacatgccacgc	gtttatcaccactgctacccctcatca	(tc)17, (ca)3, (tc)5	25	292	MSG0407	
MSE0213	LG14	C.M.S.	atctgtctgcaacaaccacaccac	gagtatgtaggtcgtccacacccc	(ag)4, (ta)6, (ag)5, (ac)6, (ac)3	6	221	AB623921	
MSG0812	LG14	C.M.S.	acatcaaacctcgcgatcatacaa	gtttcgtccatctctgtcattttctcg	(ag)15	15	163	AB624247	
MSG0425	LG14	SAK	agacaactttgttggaggagatga	gtttcttttgatcgtcgaatttgaaggg	(ag)11, (ag)13	24	128	AB624064	
MSE0054	LG14	SAK	caacaagaacgaaacaactctccg	gagagctggaatgggtgtctgagt	(ta)4, (tc)15	15	184	FS947912	Chapter 2
MSG0689	LG14	SAK	atattggaagcaaaccacaagtgc	gtttaaccgtcggtaggactcgatg	(ag)15, (ag)3	18	222	AB624195	
MSG0121b	LG14	SAK	gagcaaagaagttgtgatttcgcc	attttgagatttcgaagctccacc	(ag)11	11	264	AB623989	
MSG0645	LG14	SAK	atcgctcgagtagactgcgtgtat	gtttaagaacgacgaaacccactttga	(tc)19	19	268	AB624169	
MSE0251	LG14	SAK	cttcactccctcccatcaagcact	cacgaaagcagcaaaggtgactaa	(tc)12	12	246	AB623938	
MSE0021	LG14	SAK	tccatccatccataccattcttca	tatggaatgcatccaacaacaagg	(tc)20, (ta)10, (ta)3	20	285	FS951578	Chapter 2
MSG0519	LG14	SAK	atgttcctcctccttaccttgtcg	gtttccaagctcacaacagaaccagc	(tc)22	22	167	AB624108	
MSG0522	LG14	Kana-Ck17	atttttccagatatttgttgccgc	gtttcagcccaaaggtgagagtgttt	(tc)5, (tc)14, (ac)3	22	286	AB624109	
MSG0539	LG14	Kana-Ck17	ataaaacggggtgtctacacttgg	gtttcgagtggtctgttcaaaatctgc	(ag)21	21	206	AB624122	
MSG0581	LG15	C.M.S.	agttagactgatccacgcatctga	gtttatagacgttcggccaacaagaca	(ag)20	20	282	AB624140	
MSG0755	LG15	C.M.S.	attgaagtggtagagggtcacggt	gtttacaaactcgtcactttcaaaccca	(ag)17	17	217	AB624223	
MSG0458	LG15	C.M.S.	agaggtgtctacagagatcgtcgc	gtttcgaaaattatccccgtaggacc	(ag)14, (ga)3, (ag)3, (ta)3	23	272	AB624083	
MSG0699	LG15	C.M.S.	atgcgacagtgttgctgagatttt	gtttcaaaaatggggtgtctacagaggg	(ag)18	18	249	AB624199	
MSG0344	LG15	C.M.S.	atgtagtgagcccaatgagcaaga	gtttcccacgtatatggttggatgt	(ag)23, (ag)4	27	257	MSG0344	
MSG0481	LG15	SAK	attaggcaacctttctgccacctt	gtttacagcagcaacagcaacaacag	(tc)4, (tc)8, (ta)5, (ct)3, (tg)3	23	274	AB624096	
MSG0389	LG15	SAK	agaagcaacgaagatccactgaag	gtttctatccttcgtggcaatggattt	(ag)22, (ag)3	25	292	MSG0389	
MSG0512	LG15	Kana-Ck17	atgatgtagcacctccaatctcca	gtttagcctcgttggccacaaactatt	(ac)3, (tc)19	22	216	AB624106	
MSG0804	LG15	Kana-Ck17	atgccagaagggaagatttagaa	gtttaagtgggttgatttggagggaaat	(ag)13, (gt)3	16	120	AB624242	

Table 3.2. CAPS and STS markers on the reference map of tea.

Marker name	Linkage group	Mappable material[1)]	Marker system	Forward primer sequence	Reverse primer sequence	Expected Amplicon size (bp)	Restriction Enzymes	Source (Accession No.)	Reference
PAL	LG03	SAK	CAPS	tccatcaatctatacacctacctg	ccttctttggtcctcctatgtga	497	*Hpa*II	D26596	Kaundun and Matsumoto (2002)
GS1:1	LG03	SAK	CAPS	tttgaggtcatcaaaaaggccatt	acacaaaggttgggttgggattatg	919	*Alw*26I, *Rsa*I	AB115183	Ujihara *et. al.* (2011)
REM014	LG03	SAK	STS (SNP)					AB623062, AB623063	
AMT1:1	LG04	SAK	CAPS	ccagcgcttccaacacaaacc	ccttctgaaagtggtttgaacagagc	1358	*Dde*I, *Rsa*I	AB117640	Ujihara *et. al.* (2011)
OpT_18_200STS	LG05	C.M.S.	STS (SNP)					AB623060, AB623061	
PR1Like	LG08	C.M.S.	CAPS	gtccaccttctcacaaaatttccctataaa	ctctccatgacaacatacatggctttttaa	707	*Hha*I	AB015047	Tomimoto *et al.* (1999)
AMT1:2	LG09	SAK	CAPS	acctccggcgccctcttgtt	gctgatacgtaggaagtttgactctg	1014	*Alw*26I	AB114913	Ujihara *et. al.* (2011)
PAL_Psd1[2)]	LG12	Kana-Ck17	CAPS	cagaactccaattccttgcc	cttaagacaaggactgaaaaacattgttcc	804	*Hha*I		
GS1:2	LG15	C.M.S.	CAPS	tggtgaaggcaacgagcgtc	aaagatgataataaaattaaccccaccatattgt	1140	*Hin*fI, *Rsa*I	AB115184	Ujihara *et. al.* (2011)

1) C.M.S. means core marker set.

2) Though the PCR primer set for CAPS marker PAL_Psd1 was designed to search polymorphism of PAL, the primer set generate a CAPS mapped another locus.

Table 3.3. Summary of the linkage groups on the tea (*Camellia sinensis*) reference map.

Linkage group	Genome	Genetic distance (cM)	Total distance (cM)[a]	Number of loci	SSR loci	RAPD loci[b]	CAPS and STS loci
LG 1	Core	110	120	18	18		
	Sayamakaori	120		43	21	22	
	Kana-Ck17	108		44	27	17	
LG 2	Core	104	106	17	17		
	Sayamakaori	102		32	17	15	
	Kana-Ck17	106		40	22	18	
LG 3	Core	97	100	23	23		
	Sayamakaori	101		62	31	28	3
	Kana-Ck17	111		60	32	28	
LG 4	Core	98	98	21	21		
	Sayamakaori	108		52	26	25	1
	Kana-Ck17	72		55	28	27	
LG 5	Core	67	97	13	12		1
	Sayamakaori	94		37	15	21	1
	Kana-Ck17	86		38	16	26	1
LG 6	Core	86	92	31	31		
	Sayamakaori	92		85	45	40	
	Kana-Ck17	88		75	34	41	
LG 7	Core	84	91	14	14		
	Sayamakaori	92		24	16	8	
	Kana-Ck17	86		31	19	12	
LG 8	Core	90	90	31	30		1
	Sayamakaori	96		69	41	27	1
	Kana-Ck17	97		59	39	19	1
LG 9	Core	85	87	23	23		
	Sayamakaori	71		52	32	19	1
	Kana-Ck17	97		60	30	30	
LG 10	Core	78	84	16	16		
	Sayamakaori	86		43	21	22	
	Kana-Ck17	90		45	19	26	
LG 11	Core	74	79	15	15		
	Sayamakaori	77		29	19	10	
	Kana-Ck17	77		35	22	13	
LG 12	Core	60	75	16	16		
	Sayamakaori	58		39	23	16	
	Kana-Ck17	80		48	21	26	1
LG 13	Core	62	73	15	15		
	Sayamakaori	65		47	21	26	
	Kana-Ck17	70		38	16	22	
LG 14	Core	62	64	20	20		
	Sayamakaori	85		62	28	34	
	Kana-Ck17	58		47	22	25	
LG 15	Core	61	61	6	5		1
	Sayamakaori	57		25	7	17	1
	Kana-Ck17	73		26	7	18	1
Total	Core	1218	1317	279	276	0	3
	Sayamakaori	1305		701	363	330	8
	Kana-Ck17	1298		701	354	348	4
Total marker number without duplication				1123	441	678	9

[a] Total distance calculated from the core LG plus the extra region of the parent maps.
[b] Not including RAPD-STS.

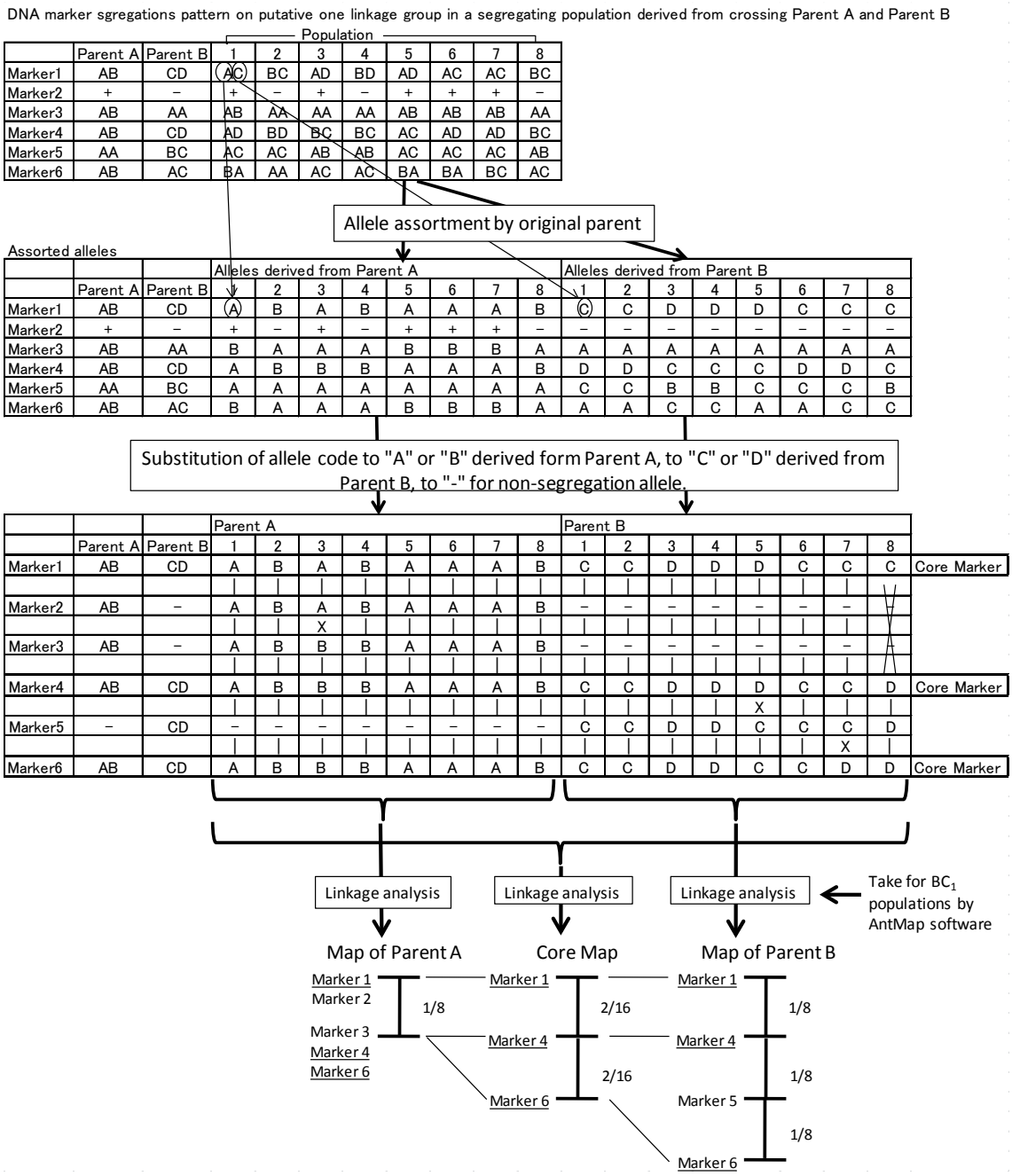

Fig. 3.1. Pattern diagram of the construction of the three-line reference map based on pseudo-testcross theory.

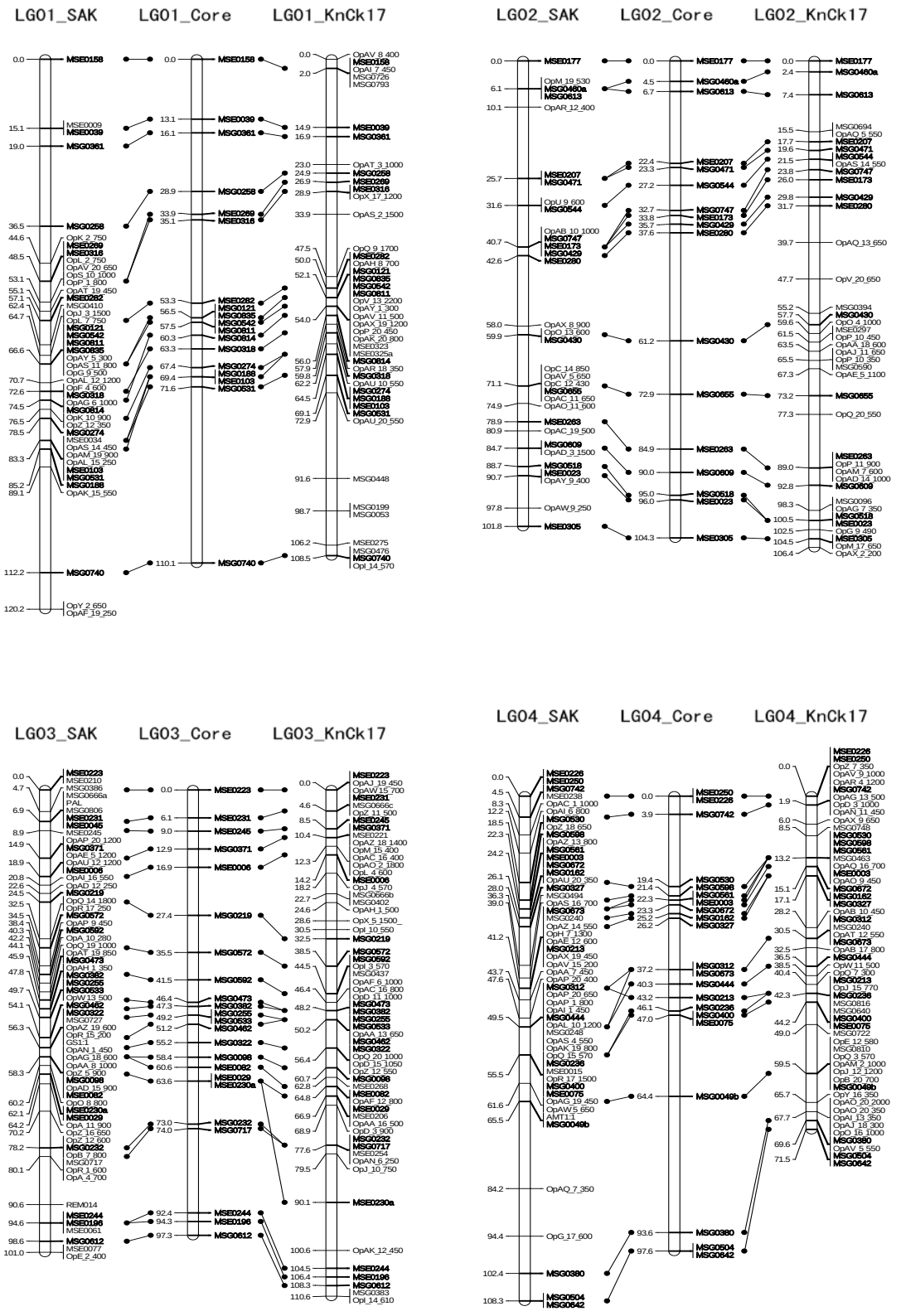

Fig. 3.2. Three-line reference genetic linkage map of tea (*Camellia sinensis*) developed by using an F_1 population derived from reciprocal crosses between 'Sayamakaori' (SAK) and Kana-Ck17 (KnCk). Locus names are given to the right of each map; genetic distances (in cM) are to the left. Core markers are shown in bold.

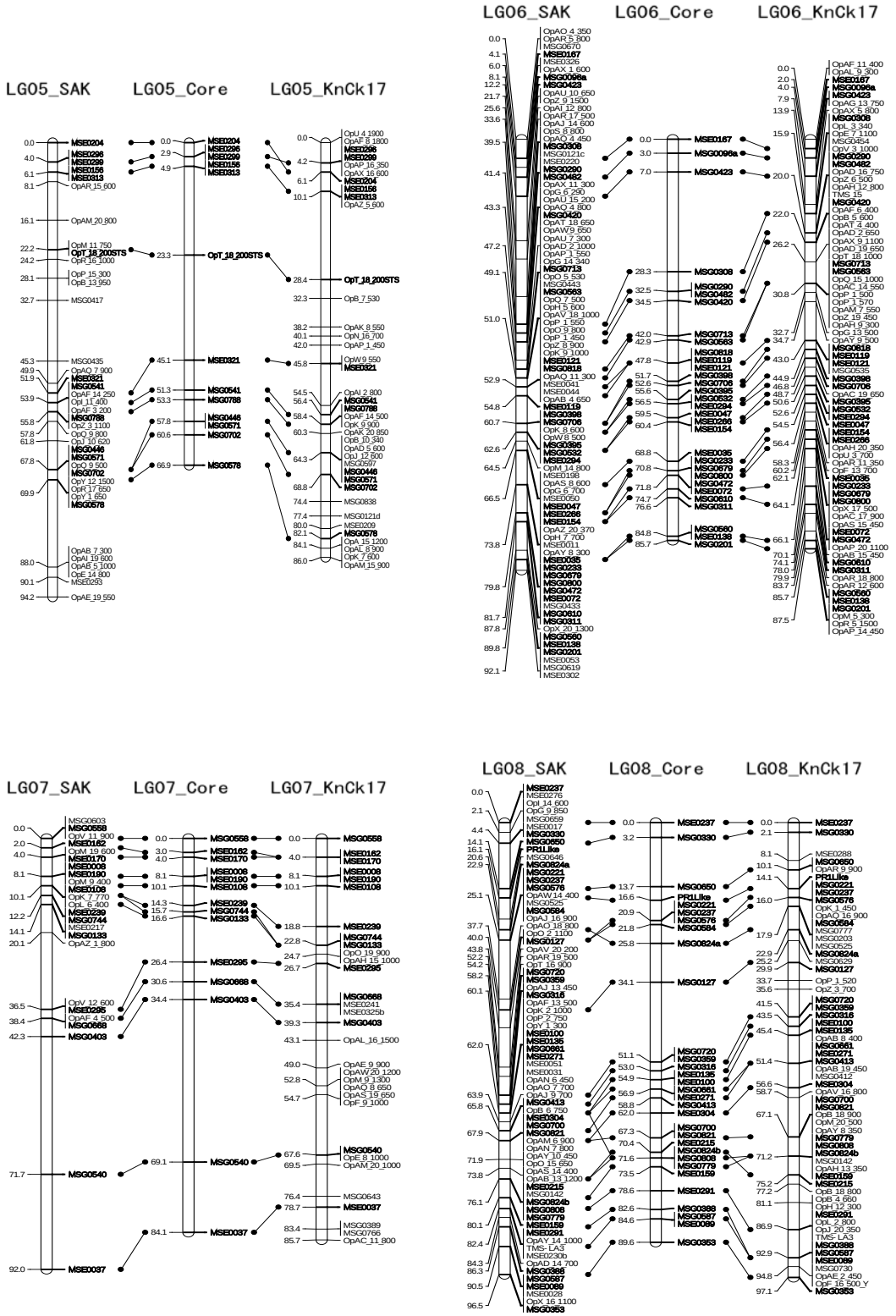

Fig. 3.2. (continued)

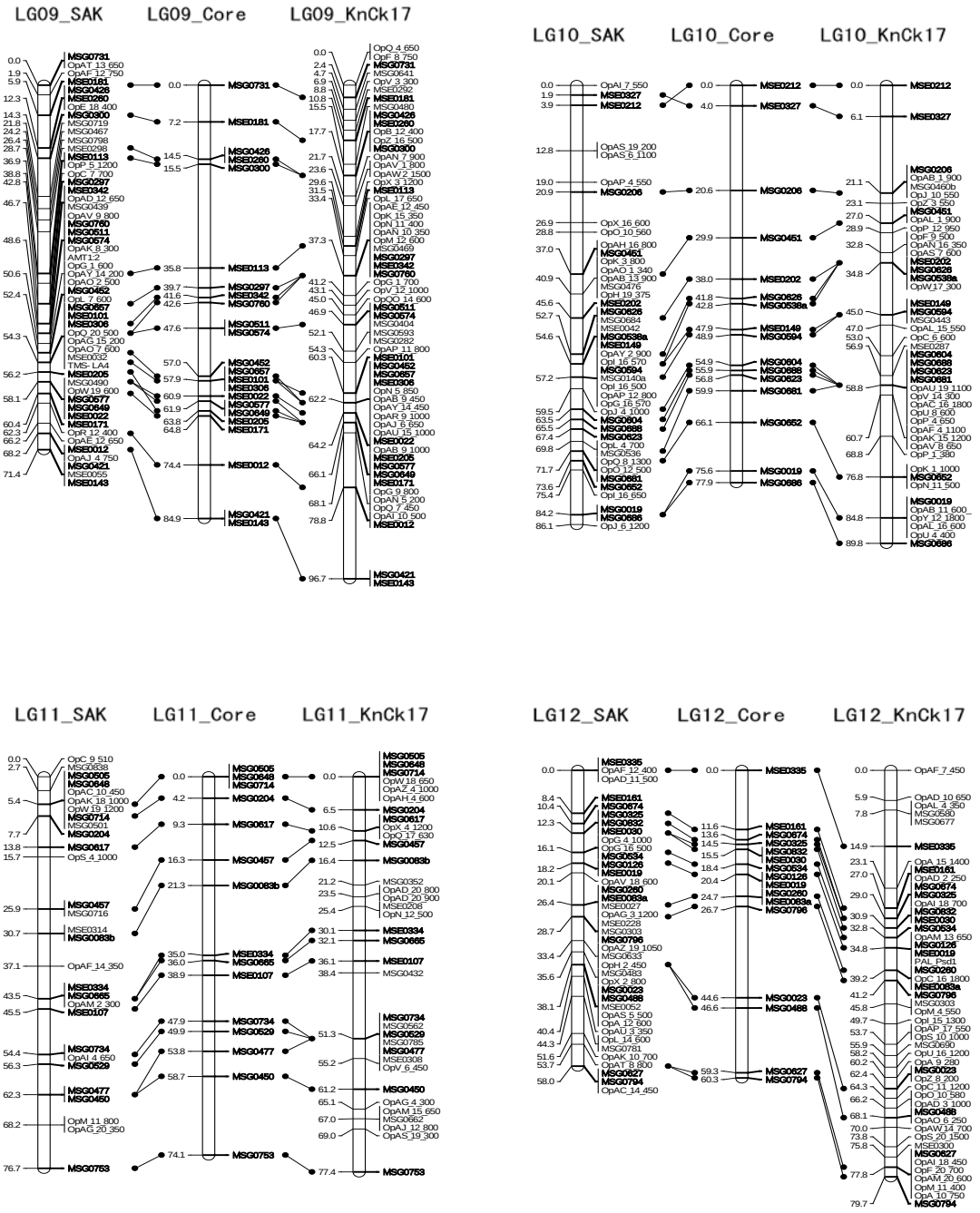

Fig. 3.2. (continued)

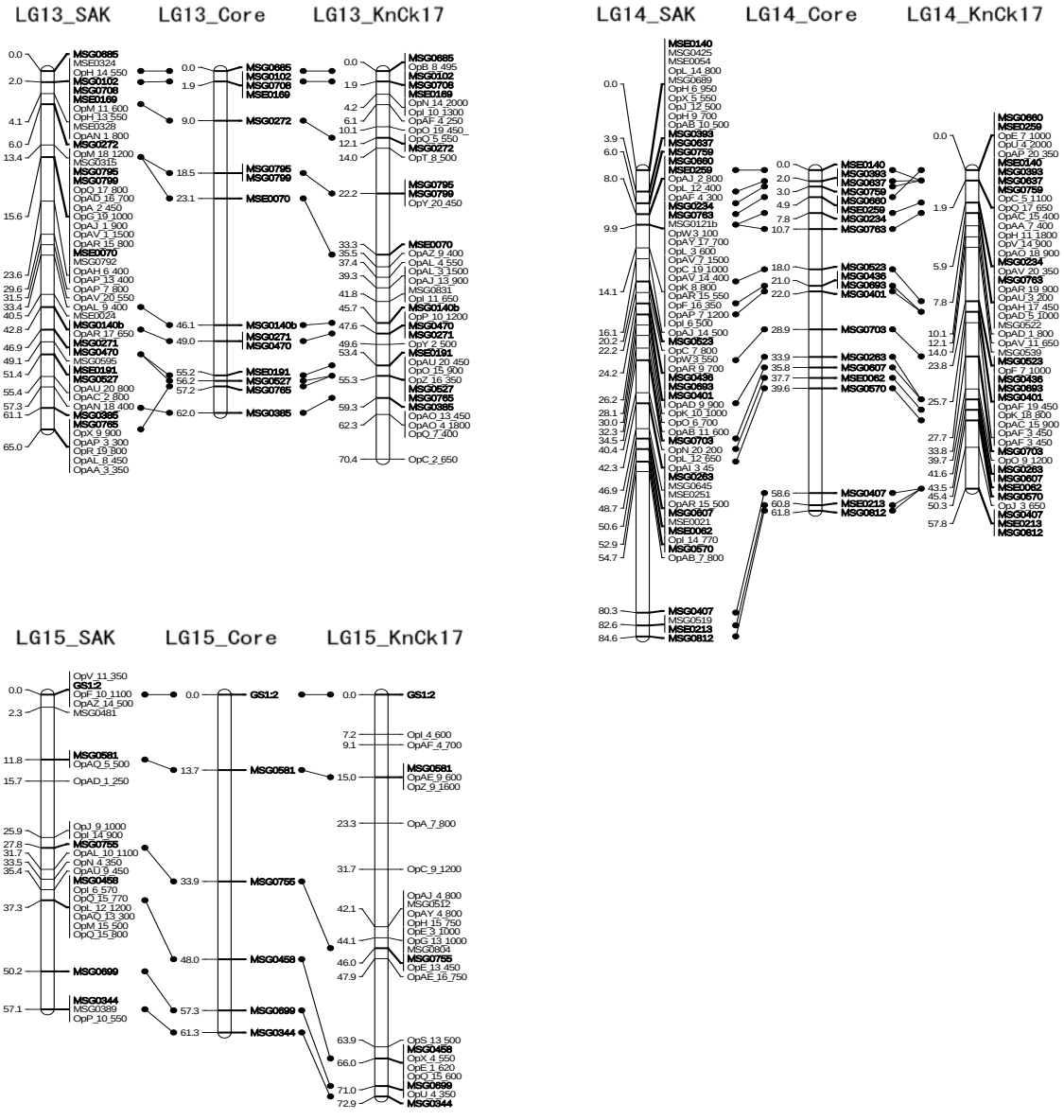

Fig. 3.2. (continued)

CHAPTER 4 Worldwide core collections of tea based on genome-wide SSR markers

4.1 Introduction

The center of diversity of *C. sinensis* is considered to be southwestern China and surrounding areas (Hashimoto and Takashi 1978, Hashimoto 2001), where the genetic diversity is highest (Yao et al. 2012). Tea plants grown here are thought to have been domesticated directly from wild ancestors (Li 1983). All tea plants grown elsewhere originated in China or India (Gunasekare 2012, Kamunya et al. 2012, Tanaka 2012, Ercisli 2012).

In spite of the very long history of the use of tea, modern breeding has a short history. Because tea is a tree crop, the generation time (duration from seed to seed) is usually 4 to 5 years, so it is difficult to shorten the breeding cycle. For this reason, the degree of genetic improvement from wild tea plants to elite cultivars is much less than in major herbaceous crops. In addition, the most important trait in tea breeding is drinking quality. Because the standards for the evaluation of tea quality are different for each of the several types of tea and differ among countries, tea breeders select breeding parents based on country-specific evaluation standards. A few elite cultivars are used as breeding materials, so the genetic diversity of breeding populations is narrow. Because high-quality progeny generally result from crossing high-quality parents, breeders tend to give the highest priority to such crossing combinations. For instance, in Japan, the green tea cultivar ‘Yabukita’ occupies about 75% of tea fields and is frequently used as breeding material. Therefore, many other major cultivars are derived from ‘Yabukita’ (Tanaka 2012). Although it is necessary to increase the genetic diversity of breeding populations in order to advance tea breeding, there is not enough information about which germplasms are useful for different breeding programs.

To acquire such information, I need to assess genetic diversity. Many DNA marker analyses of genetic diversity have already been reported. Several studies of various germplasms of both var. *sinensis* and var. *assamica* from multiple countries have revealed that collections from China and India have high levels of genetic diversity, but

these studies analyzed only a few dozen accessions (Wachira et al. 2001; Kaundun and Matsumoto 2002, 2003; Yao et al. 2008). More recently, some studies analyzed several hundred germplasms using SSR or AFLP markers (Ohsako et al. 2008, Yao et al. 2012, Raina et al. 2012, Fang et al. 2012), but only domestic germplasms from Japan, China, or India.

For the efficient mining of breeding materials with targeted phenotypes, core collections, which represent as much as possible of the full range of genetic diversity with a minimum number of accessions, are useful resources for breeders. Two studies that developed core collections of tea have so far been reported (Wang et al. 2011, Raina et al. 2012). Using phenotypic data, Wang et al. (2011) selected a core collection of 532 accessions from 2557 Chinese accessions and 108 accessions from other countries. From 1664 accessions and clones of Indian hybrid tea, Raina et al. (2012) selected a core collection of 105 accessions, 52 of which were selected by using AFLP genotype data and 53 by phenotypic data. However, until now, no core collection of tea has been selected from worldwide genetic resources on the basis of genotype only.

NIVTS has a worldwide tea germplasm collection of more than 7800 accessions obtained from 14 countries or regions over the past century (Takeda 2002). A large-scale phenotypic analysis of this collection (Takeda 2002) showed it to have wide diversity in various traits. Thus, this collection is a suitable resource for the selection of worldwide core collections.

4.2 Materials and methods

4.2.1 Plant materials

From the 7800 accessions of *C. sinensis* germplasm at NIVTS, I selected 788 accessions for SSR marker genotyping based on a consideration of the passport data (including country of origin and collection site), the biological status (e.g., wild, landrace, breeder's line), and previous knowledge about genetic diversity (Takeda 2002, Matsumoto et al. 2002), to cover as much variation as possible. This set originates from all 14 origins from which NIVTS tea germplasms were collected (Fig. 4.1; Table 4.1), but biased toward China and India, where the genetic diversity is likely to be highest.

4.2.2 DNA extraction and SSR marker analysis

Total DNA of each accession was extracted from fresh leaves by a method using diatomaceous earth and a spin filter (Tanaka and Ikeda 2002). As SSR markers for genotyping, I selected 23 loci from the core map of *C. sinensis* (Chapter 3) to cover the whole genome evenly (Tables 4.2, 4.3). PCR was performed in a 10-μL reaction mix including 20 ng of total DNA, 10× PCR Gold buffer (Life Technologies, Carlsbad, CA, USA), 0.8 μL of 8 mM dNTPs, 0.1 U of AmpliTaq Gold Polymerase (Life Technologies), 0.8 μL of 25 mM $MgCl_2$, and 1 μM each of the forward and reverse primers. The PCR reactions were carried out in a GeneAmp 9700 thermal cycler (Life Technologies) according to the following touchdown PCR cycling program: 95 °C for 5 min; 95 °C for 1 min, 62 °C for 30 s, and 72 °C for 1 min; 13 cycles at annealing temperatures decreasing by 0.5 °C per cycle; 25 cycles of 95 °C for 1 min, 55 °C for 30 s, and 72 °C for 1 min; and a final extension at 72 °C for 10 min. PCR products were directly fluorescently labeled with R110-ddUTP by the single-tube method (Inazuka et al. 1996). The labeled PCR products were analyzed with an ABI Prism 3130*xl* Genetic Analyzer (Life Technologies), and the size of the fragments was determined with the GeneMapper v. 4.0 software (Life Technologies).

4.2.3 Population analysis

Using the genotype data from the SSR marker loci, I calculated expected heterozygosity (H_E), observed heterozygosity (H_O), and polymorphic information content (PIC) in the PowerMarker software (Liu and Muse 2005). Since the numbers of accessions differed among origins, I calculated allelic richness by the rarefaction method in the FSTAT 2.9.3.2 software (Goudet 1995) to allow comparison.

To clarify the population structure, I performed four analyses. First, I investigated the population structure using the Bayesian clustering algorithm implemented in the Structure software (Pritchard et al. 2000). This analysis evaluated 1 to 10 population genetic clusters (K) with 10 runs per K value. For each run, the initial burn-in period was set to 50,000, with 50,000 Markov-chain Monte Carlo iterations. The most likely number of clusters was determined by using the DeltaK method (Evanno et al. 2005) and the coefficients of correlations among the 10 runs were determined in the CorrSieve

software (Campana et al. 2011). Bar plots of the Structure results were drawn in the CLUMPP software (Jakobsson and Rosenberg 2007) and the Distruct software (Rosenberg 2004).

I performed AMOVA (Excoffier et al. 1992) in the GenAlEx v. 6.5 software (Peakall and Smouse 2012) to elucidate the extent of genetic variation among and within populations, and calculated pairwise F_{st} (Cockerham and Weir 1993), treating accessions from each origin as a population. Statistical significance was determined from 1000 permutations.

The populations were grouped by means of principal-coordinates analysis (PCOA) of Nei's genetic distance matrix (D_s: Nei 1978) in the GenAlEx software.

The shared allele distance (Chakraborty and Jin 1993) was calculated, and a phylogenetic tree was constructed by the neighbor-joining method in the Populations v. 1.2.32 software (Langella 1999).

4.2.4 Selection for core collections

Using the coverage of allele numbers at the 23 SSR marker loci to evaluate the genetic diversity, I selected core collections in PowerMarker. This software uses a simulated annealing algorithm to capture the maximum number of alleles in a given number of accessions. I selected four core collections (24, 48, 96, and 192 accessions). For each number of accessions, 10 replicates of selection of accessions included in the core collections were used. First, I selected the Core 24 set from all 788 accessions to hold 'Yabukita', 'Sayamakaori', 'Minamisayaka', 'Sizu Inzatsu 131', and 'Kan-Ck17', which are the parents of populations used in linkage-map construction and genetic analysis of important agronomic traits (Tanaka and Sawai 2005; Taniguchi et al. 2010, Chapter 3). Next, I selected Core 48 to contain all accessions of Core 24 plus an additional 24 accessions. I similarly selected Core 96 and Core 192.

4.2.5 Phenotyping of germplasms

To verify whether the core collections covered the range of phenotypic diversity within the set of 788 accessions, I investigated floral morphology and the chemical composition of first-flush young shoots. Floral morphology is useful for the analysis of

genetic relationships because of its high heritability, and has been used to evaluate relationships among genetic backgrounds in *C. sinensis* (Takeda and Toyao 1980, Toyao and Takeda 1999, Chen J et al. 2005). I evaluated each accession for the density of the pubescence of the ovary, level of style protrusion, number of split style segments, depth of style splitting, and extent of style bending.

The chemical composition of first-flush young shoots is important for tea quality. I harvested first-flush young shoots, steamed them for 37 s, and dried them at 70 °C for 4 h. Dry-milled powder of the shoots was analyzed for the contents of total N, total free amino acids, theanine, tannin, and caffeine by near-infrared reflectance analysis with a tea component analyzer (GT-8S; Shizuoka Seiki, Fukuori, Japan).

4.2.6 Statistical analyses

I compared the CVs of the chemical constituents of young leaves, and performed *F*-tests for equality of variance between all accessions and all core collections.

4.3 Results

4.3.1 SSR genotyping and diversity statistics

Among the 788 tea germplasms, allele numbers averaged 26.9 (range, 14–42), with H_E = 0.86 ± 0.014 (mean ± SEM), H_O = 0.62 ± 0.035, and PIC = 0.85 ± 0.015 (Table 4.3). As expected, accessions from China, Taiwan, India, and Sri Lanka had high heterozygosity values, and those from Japan had low values (Table 4.4). Allelic richness showed the same tendencies.

4.3.2 Population structure analysis

Population structure analysis showed that DeltaK reached its maximum at K = 2 (Fig. 4.2). Correlation coefficients were >0.99 for K = 2 and 3, and an additional peak appeared at K = 8. The DeltaK method of Evanno et al. (2005) often results in K = 2, owing to markedly low likelihoods for K = 1 in all analyses (Vigouroux et al. 2008). Therefore, I regarded K = 2, 3 and 8 as appropriate results. At K = 2, cluster 1 favored accessions from Japan and cluster 2 favored exotic germplasms (Fig. 4.3). At K = 3,

cluster 1 was retained, but cluster 2 was divided into cluster 2-1, comprising China and part of India, and cluster 2-2, comprising Southeast Asia, Bangladesh, part of India, and Sri Lanka. Given the origins of the accessions in these subclusters, cluster 2-1 represents var. *sinensis* and cluster 2-2 var. *assamica*. At $K = 8$, cluster 2-1 at $K = 3$ was divided into four clusters, and cluster 2-2 into three clusters: clusters 2-11 (accessions from western China), 2-12 (Korea), 2-13 (southwestern China), and 2-14 (Darjeeling, India); and clusters 2-21 (India), 2-22 (Taiwan), and 2-23 (Vietnam).

4.3.3 AMOVA

The genetic variation was 14% among populations and 86% within populations (Table 4.5). I calculated pairwise F_{st} values for all countries except Thailand, from which only one accession originated. F_{st} ranged from 0.008 to 0.217 (Table 4.6). Most F_{st} values between Japan and other countries (in particular, Southeast Asia and Bangladesh) were high, indicating that tea genetic resources from Japan are highly differentiated from the rest.

4.3.4 Principal-coordinates analysis

PCO1 and PCO2 explained 40.0% and 20.1%, respectively, of the variance in the genotype data (Fig. 4.4). In the scatterplot, the Japanese accessions (PCO1 > 0) were clearly separated from the exotic accessions (PCO1 < 0). Accessions from Korea and China (where var. *sinensis* was dominant) clustered in PCO2 > 0, and those from Southeast Asia and Sri Lanka (where var. *assamica* was dominant) in PCO2 < 0.

4.3.5 Phylogenetic tree

Phylogenetic analysis placed most accessions from Japan in one clade (Fig. 4.5). The inclusion of var. *assamica* accessions in this clade was unexpected. Exotic accessions were divided into var. *sinensis* and var. *assamica*. Although most of the accessions from China were placed in the exotic var. *sinensis* group, some were placed in the Japanese clade. Accessions from India were widely distributed in both the var. *sinensis* and the var. *assamica* clades. Two sub-clades within the same parent clade consisted only of wild-growing tea plants from Taiwan and Vietnam. Accessions from

Sri Lanka were placed with accessions from India. Accessions from Iran and Georgia were placed with accessions from China. These results are consistent with the history of the expansion of tea cultivation.

4.3.6 Selection and evaluation of core collections

I selected the Core 192, Core 96, Core 48, and Core 24 collections to cover as many alleles as possible (Tables 4.4, 4.7). The set of 788 accessions had 619 alleles. Core 192 had 616 alleles (99.5%, Table 4.4), followed by Core 96 (546 alleles, 88.2%), Core 48 (462, 74.6%), and Core 24 (368, 59.5%). Core 192 and Core 96 covered most of the variation in floral morphology and chemical constituents of first-flush young leaves (Figs. 4.6, 4.7). In several phenotypes, Core 24 and Core 48 lacked a maximum or minimum class, but in most traits they represented the total variation in spite of their small sizes.

Core 192 and Core 96 showed almost equal CVs for all components; Core 48 and Core 24 showed smaller CVs in several components (Table 4.8). Although the variances of theanine in all core collections were significantly larger than in the full set, no components within the core collections showed significantly smaller variances than the full set (Table 4.8).

4.4 Discussion

4.4.1 Genetic diversity and population structure of the tea germplasm collection

I analyzed a total of 788 accessions of var. *sinensis* and var. *assamica* from 14 origins to reveal the genetic diversity of tea genetic resources.

Germplasms from Japan had lower genetic diversity than those from other countries (Fig. 4.4; Table 4.4), and germplasms from Taiwan, China, India, and Sri Lanka had higher diversity. Japan, Taiwan, and Sri Lanka are all islands, yet only germplasms from Japan had low genetic diversity. Most of the cultivated clones in Taiwan were introduced from China and India (Lai et al. 2001). These origins would explain the high genetic diversity of Taiwanese tea. Almost all tea plants grown in Sri Lanka have been introduced from different parts of India since the 19th century, also with high genetic diversity, and morphologic traits point to both var. *assamica* and var. *sinensis*

(Gunasekare 2012). The results of our Structure analysis and PCOA support this idea, and indicate that some accessions are hybrids between the two varieties. In contrast, tea seeds were brought to Japan from only a limited origin in China, and the founder effect therefore explains the narrow genetic base.

The low molecular variance among populations (14%) means limited differentiation of populations, indicating that tea plants from the 14 origins have not been exposed to strong selection and genetic drift and that accessions from some countries, notably India, had been introduced from other regions by humans.

As *C. sinensis* is classified into two major varieties, it would be expected that population structure analysis would identify two corresponding groups. However, our analysis clearly divided the accessions between Japanese and exotic accessions. Considering that most Japanese accessions are var. *sinensis*, this population structure seems strange. Two reasons might explain this result.

The first is the founder effect that occurred when tea seeds were brought from China to Japan. A dramatic change in the composition of genotypes in Japan is consistent with this origin (Matsumoto et al. 2002).

The second possible reason is a bias in the SSR markers that I used. I selected the markers from a tea reference map (Chapter 3) constructed by using an F_1 population derived from a Japanese cultivar, 'Sayamakaori', and a Chinese line, Kana-Ck17. Therefore, the markers were polymorphic between the Japanese and Chinese germplasms, which are both var. *sinensis*. The resultant marker-selection bias was strong enough to suppress the division of the exotic accessions into var. *sinensis* and var. *assamica* when $K = 2$.

As stated above, the DeltaK method often results in $K = 2$ because of markedly low likelihoods for $K = 1$. Marker-selection bias may also favor $K = 2$ in the DeltaK method. Our results show that in selecting DNA markers for genetic diversity and population structure analysis, I should not limit our selection only to markers with polymorphisms between genotypes with specific genetic backgrounds.

4.4.2 Worldwide core collections of tea genetic resources

For active use of genetic resources in breeding, core collections (which represent

the genetic diversity of the whole collection with minimum sets) play important roles. To date, two core collections of tea had been reported (Wang et al. 2011, Raina et al. 2012). The Chinese core collection of Wang et al. (2011) was selected by using phenotypic data. Although 4.0% of the initial collection were accessions introduced from other countries, 14.3% of the core collection were introduced. This means that the introduced accessions have many alleles not present in Chinese domestic accessions. Because only 4.0% of the initial accessions were introduced, it is difficult to conclude that the core collection covers the genetic diversity of worldwide tea germplasms. In addition, the number of accessions in the core collection, 532, is too large for detailed experiments. The core collection of Raina et al. (2012) was selected from a germplasm collection of Indian hybrid teas. Although some tea plants cultivated in India originate from China, Raina et al. (2012) did not analyze germplasms from other tea-producing countries, so the Indian core collection is also not a worldwide core collection. In addition, they selected core collections by genotype and phenotype independently, but did not consider the redundancy between the two groups. In contrast, our core collections cover the worldwide diversity of tea and are of convenient sizes for experiments.

I selected core collections with 192, 96, 48, and 24 accessions from the 788 accessions and evaluated their coverage of phenotypic diversity. In spite of the marker-selection bias, the Core 96 and Core 192 sets covered almost the entire range of phenotypic values tested, and will therefore be useful for screening of novel breeding materials or for genome-wide association studies of important agronomic traits. Although the Core 48 and Core 24 sets covered fewer alleles, the variances in the chemical components of first-flush young leaves were not significantly smaller than those of the full set. Therefore, these two core sets will be useful for some genetic diversity research, such as polymorphism testing in the development of novel DNA markers or experiments for evaluating the genetic variance underlying phenotypic values, for which phenotyping methods are very laborious.

I found a division between Japanese and exotic accessions, the latter of which were divided between var. *sinensis* and var. *assamica*. The resultant information about the genetic diversity of accessions from 14 origins will be useful for planning the exploration, use, and conservation of genetic resources. The four core collections I

selected will enhance research on the genetic diversity of *C. sinensis*. In particular, the identification of loci underlying important agronomic traits by genome-wide association analysis will be made possible by the development of several tens of thousands of genome-wide SNPs for large-scale genotyping of core collections.

Table 4.1. List of tea germplasms used in this study

No.	Accession name	Origin	Accession No.	No.	Accession name	Origin	Accession No.
1	Yabukita	Japan	27027257	77	Mak Zai 30-5	Japan	27004473
2	Sayamakaori	Japan	27029293	78	Mak Zai 30-13	Japan	27004476
3	Minamisayaka	Japan		79	Mak Zai 88-3	Japan	27004492
4	Shiz Inzatsu 131	Japan		80	Mak Zai 88-39	Japan	27004524
5	Sayamamidori	Japan	27027252	81	Mak Zai 88-24	Japan	27004509
6	Asatsuyu	Japan	27027248	82	Mak Zai 88-19	Japan	27004504
7	Z 1	Japan		83	Mak Zai 88-15	Japan	27004500
8	Benifuki	Japan		84	Asahi	Japan	27029294
9	Miyamakaori	Japan		85	Komakage	Japan	27029291
10	Shizu Zai 16	Japan		86	Kyomidori	Japan	
11	Iwate 3	Japan		87	Okunoyama	Japan	
12	Iwate 15	Japan		88	Kna Zai Kyo 88-4	Japan	27013059
13	Iwate 23	Japan		89	Kna Zai Kyo 88-11	Japan	27013066
14	Akita 38	Japan		90	Kna Zai Kyo 88-34	Japan	27013089
15	Akita 17	Japan		91	Kna Zai Kyo 88-71	Japan	27013126
16	Akita 7	Japan		92	Kna Zai Kyo 88-80	Japan	27013173
17	Akita 28	Japan		93	Kna Zai Kyo 88-108	Japan	27013201
18	Mak Zai 1-2	Japan	27004129	94	Kna Zai Kyo 88-117	Japan	27013209
19	NG 41	Japan	27002122	95	Kna Zai Kyo 88-135	Japan	27017592
20	Mak Zai 3-2	Japan	27004148	96	Kna Zai Kyo 88-147	Japan	27017603
21	Mak Zai 3-6	Japan	27004150	97	Kna Zai Kyo 88-189	Japan	27017645
22	Mak Zai 10-3	Japan		98	Kna Zai Kyo 88-210	Japan	27017666
23	Mak Zai 11-17	Japan		99	Kna Zai Kyo 88-221	Japan	
24	Mak Zai 8-11	Japan	27004162	100	Kna Zai Kyo 88-227	Japan	
25	Mak Zai 13-2	Japan	27004174	101	Kna Zai Kyo 88-237	Japan	
26	Mak Zai 9-1	Japan	27004163	102	Kna Zai Kyo 88-251	Japan	
27	Mak Zai 171-2	Japan		103	Na Zai Kyo 2	Japan	27017676
28	Mak Zai 171-3	Japan		104	Na Zai Kyo 37	Japan	27017711
29	K 3	Japan	27002097	105	Na Zai Kyo 49	Japan	27017723
30	Hachioji Koro	Japan		106	Na Zai Kyo 64	Japan	27017738
31	Niigata 13	Japan		107	Na Zai Kyo 88	Japan	27017762
32	Niigata 29	Japan		108	Na Zai Kyo 191	Japan	27017865
33	Niigata 66	Japan		109	Mak Zai 80-30	Japan	27004329
34	Mak Zai 116-10	Japan	27004143	110	Mak Zai 148-13	Japan	27005147
35	Mak Zai 2-4	Japan	27004132	111	Mak Zai 148-27	Japan	27005161
36	Mak Zai 93-8	Japan	27004251	112	Mak Zai 147-16	Japan	27005120
37	Mak Zai 92-9	Japan	27004242	113	Mak Zai 147-30	Japan	27005134
38	Mak 73-2	Japan	27004236	114	Mak Zai 81-1	Japan	
39	Kna Zai No 30	Japan	27002135	115	Mak Zai 81-24	Japan	27004354
40	Gifu Zai 30-8	Japan		116	Mak Zai 145-2	Japan	27005046
41	Gifu Zai 30-2	Japan		117	Mak Zai 151-8	Japan	27005193
42	Gifu Zai 30-11	Japan		118	Mak Zai 84-34	Japan	27004438
43	K 1	Japan	27002095	119	Mak Zai 87-2	Japan	27004445
44	Kna Zai Sz 20	Japan	27029320	120	Mak Zai 82-22	Japan	27004376
45	Siz 7132	Japan		121	Mak Zai 82-8	Japan	27004365
46	Kuritawase	Japan		122	Mak Zai 83-38	Japan	
47	NG 10	Japan	27002112	123	Mak Zai 113-3	Japan	27004530
48	Mak Zai 20-4	Japan	27004182	124	Mak Zai 114-1	Japan	27004534
49	Sizu At 5	Japan		125	Mak Zai 109-2	Japan	27004527
50	Sizu At 10	Japan		126	Yamatomidori	Japan	27027259
51	Sizu At 11	Japan		127	Mak Zai 127-14	Japan	27004304
52	Sizu At 14	Japan		128	Mak Zai 128-19	Japan	27004312
53	Sizu At 14	Japan		129	Mak Zai 126-13	Japan	27004287
54	Kanmuriyama Zai 24-2	Japan		130	Mak Zai 126-5	Japan	27004280
55	Mak Zai 69-13	Japan	27004229	131	Mak Zai 58-1	Japan	
56	Mak Zai 65-16	Japan	27004196	132	Mak Zai 89-2	Japan	27004253
57	Mak Zai 66-2	Japan	27004198	133	Kna Zai Shimane 5-5	Japan	
58	Mak Zai 70-4	Japan	27004230	134	Kna Zai Shimane 4-4	Japan	
59	Mie At 107	Japan		135	Kna Zai Shimane 3-1	Japan	
60	Mie At 120	Japan		136	Kna Zai Shimane 1-5	Japan	
61	Mie At 133	Japan	27004100	137	Mak Zai 29-2	Japan	27004541
62	Mie At 1012	Japan	27004102	138	Kna Zai 4	Japan	27002200
63	Mie At 1039	Japan		139	Mak Zai 72-1	Japan	27004535
64	Shiga Zai 27-14	Japan		140	Mak Zai 72-2	Japan	27004536
65	Shiga Zai 27-7	Japan		141	Mak Zai 27-1	Japan	27004537
66	Kna Zai Shig 79	Japan	27002163	142	Mak Zai 27-3	Japan	27004539
67	Kna Zai Shig 93	Japan	27002166	143	Mak Zai 27-5	Japan	27004540
68	Kna Zai Shig 27	Japan	27002146	144	Kna ZaiToku45	Japan	27002190
69	Kna Zai Shig 33	Japan	27002149	145	Kna ZaiToku137	Japan	27002195
70	Kna Zai Shig 12	Japan	27002144	146	Mak Zai 133-23	Japan	27004580
71	Kna Zai Shig 62	Japan	27002158	147	Mak Zai 132-25	Japan	27004557
72	Kna Zai Shig 70	Japan	27002161	148	Mak Zai 134-27	Japan	27004599
73	Mak Zai 32-1	Japan		149	Mak Zai 36-7	Japan	27004620
74	Mak Zai 32-3	Japan		150	Mak Zai 34-5	Japan	27004603
75	Mak Zai 33-5	Japan	27004486	151	Mak Zai 34-21	Japan	27004613
76	Mak Zai 30-1	Japan	27004472	152	Mak Zai 2769-2	Japan	27005242

Table 4.1. continued.

No.	Accession name	Origin	Accession No.	No.	Accession name	Origin	Accession No.
153	Kochi Ikegawa Yamacha 2	Japan		229	Makizono Daichaju	Japan	27005755
154	Mak Zai 2772-1	Japan	27005265	230	Mak Zai 106-4	Japan	27005010
155	KnZaiKo21	Japan	27002177	231	Mak Zai 158-3	Japan	27005031
156	Mak Zai 136-17	Japan	27004662	232	Mak Zai 162-3	Japan	27005036
157	Mak Zai 135-13	Japan	27004649	233	Mak Zai 157-4	Japan	27005018
158	Fukuoka Zai 26-15	Japan		234	Mak Zai 78-2	Japan	27005002
159	Himemidori	Japan		235	Mk Cok 18	Japan	
160	Mak Zai 3020-7	Japan	27005342	236	Mk Cok 48	Japan	
161	Mak Zai 3001-8	Japan	27005326	237	Mk Cok 349	Japan	
162	Mak Zai 3003-5	Japan	27005331	238	Mk Cok 368	Japan	
163	Mak Zai 129-7	Japan	27004695	239	K 22	Korea	
164	Mak Zai 129-14	Japan	27004699	240	K 35	Korea	
165	Mak Zai 130-20	Japan	27004724	241	Kna Korea 93-1-31	Korea	27027904
166	Mak Zai 39-18	Japan	27004687	242	Kna Korea 93-1-48	Korea	27027921
167	Mak Zai 37-6	Japan	27004679	243	Kna Korea 93-2-45	Korea	27027955
168	Mak Zai 131-26	Japan	27004735	244	Kna Korea 93-4-33	Korea	27027979
169	Saga Zai 25-24	Japan		245	Kna Korea 93-5-33	Korea	27028010
170	Saga Zai 25-14	Japan		246	Kna Korea 93-6-40	Korea	27028044
171	Mak Zai 121-4	Japan	27004742	247	Kna Korea 93-7-8	Korea	27028053
172	Mak Zai 120-4	Japan		248	Kna Korea 93-8-45	Korea	27028133
173	Mak Zai 120-2	Japan		249	Kna Korea 93-9-16	Korea	27028168
174	Mak Zai 122-12	Japan	27004760	250	Kna Korea 93-10-20	Korea	27028197
175	Mak Zai 125-7	Japan	27004767	251	Kna Korea 93-11-13	Korea	27028216
176	Mak Zai 125-1	Japan		252	Kna Korea 93-11-41	Korea	27028238
177	Mak Zai 124-4	Japan	27004766	253	Kna Korea 93-12-2	Korea	27028247
178	Mak Zai 123-4	Japan	27004762	254	Kna Korea 93-13-13	Korea	27028283
179	Kna Zai Tsushima 19-1	Japan		255	Kna Korea 93-14-1	Korea	27028300
180	Kna Zai Tsushima 20-2	Japan		256	Kna Korea 93-15-23	Korea	27028329
181	Kna Zai Tsushima 18-1	Japan		257	Kna Korea 93-16-5	Korea	27028334
182	Kna Zai Tsushima 10-2	Japan		258	Kna Korea 93-18-22	Korea	27028367
183	Kna Zai Tsushima 9-1	Japan		259	Kna Korea 93-20-17	Korea	27028395
184	Kna Zai Goto 7-2	Japan		260	Kna Korea 93-21-14	Korea	27028418
185	Kna Zai Tsushima 7-3	Japan		261	Kna Korea 93-22-5	Korea	27028440
186	Kna Zai Tsushima 6-3	Japan		262	Kna Korea 93-22-20	Korea	27028453
187	Kna Zai Tsushima 4-3	Japan		263	Kna Korea 93-23-4	Korea	27028471
188	Kna Zai Goto 9-1	Japan		264	Kna Korea 93-27-13	Korea	27028526
189	Hirado Yamacha 11	Japan	27005212	265	Kna Korea 93-28-26	Korea	27028549
190	Kna Zai Goto 4-1	Japan		266	Kna Korea 93-29-14	Korea	27028574
191	Kna Zai Goto 2-4	Japan		267	Kna Korea 93-31-2	Korea	27028586
192	Kna Zai Tsushima 1-1	Japan		268	Kna Korea 481	Korea	27002032
193	Kna Zai Tsushima 1-5	Japan		269	Mak Cok 113	Taiwan	
194	Kna Zai Tsushima 1-7	Japan		270	Mak Cok 129	Taiwan	
195	Kna Zai Tsushima 2-4	Japan		271	Ay 5306	Taiwan	27003338
196	Kna Zai No 10	Japan	27002134	272	Ay 5315	Taiwan	27003340
197	Nagasaki Zai 28-4	Japan		273	Taiwan Yamacha 25	Taiwan	27003268
198	Mak Zai 118-24	Japan	27004786	274	Taiwan Yamacha 30	Taiwan	27003273
199	Mak Zai 118-15	Japan	27004779	275	Taiwan Yamacha 37	Taiwan	27003279
200	Itsuki Yamacha 29	Japan		276	Taiwan Yamacha 40	Taiwan	27003282
201	Sagara Zai 9-2	Japan		277	Taiwan Yamacha 47	Taiwan	27003289
202	Hitoyoshi Zai 9-2	Japan		278	Taiwan Yamacha 53	Taiwan	27003295
203	Itsuki Yamacha Kozuru 32-14	Japan		279	Taiwan Yamacha 56	Taiwan	27003298
204	Mak Zai 2824-2	Japan	27005275	280	Taiwan Yamacha 62	Taiwan	27003303
205	Mak Zai 53-4	Japan	27004791	281	Taiwan Yamacha 67	Taiwan	27003307
206	Mak Zai 2948-3	Japan	27005306	282	Taiwan Yamacha 70	Taiwan	27003310
207	Mak Zai 2953-3	Japan	27005309	283	Taiwan Yamacha 73	Taiwan	27003313
208	Mak Zai 2957-4	Japan	27005315	284	Taiwan Yamacha 78	Taiwan	27003318
209	Gokase Zai 6-3	Japan		285	Taiwan Yamacha 79	Taiwan	27003319
210	Takachiho Yamacha 8-4	Japan		286	Taiwan Yamacha 84	Taiwan	27003324
211	Hinokage Zai 7-1	Japan		287	Taiwan Yamacha 88	Taiwan	27003328
212	Nishimera Yamacha 35-3	Japan		288	Miya Chinshindahpan 2	Taiwan	27025760
213	Nishimera Yamacha 14	Japan		289	Kna Ay 103	Taiwan	27013152
214	Nishimera Yamacha 36-1	Japan		290	Kna Ay 94	Taiwan	27013146
215	Shiiba Yamacha 38-10	Japan		291	Chinshindahpan	Taiwan	
216	Shiiba Yamacha 37-23	Japan		292	Taicha17	Taiwan	
217	Shiiba Yamacha 24	Japan		293	Taiwan Yamacha 12	Taiwan	27003257
218	Shiiba Yamacha 21	Japan		294	Taiwan Yamacha 18	Taiwan	27003262
219	ME 52	Japan	27025724	295	Taiwan Yamacha Nishi1	Taiwan	
220	Mak Zai 143-3	Japan	27004880	296	Taiwan Yamacha 95	Taiwan	27003335
221	Mak Zai 165-3	Japan	27004891	297	Taiwan Okinawa 9	Taiwan	27027353
222	Mak Zai 167-5	Japan	27004931	298	Taiwan Okinawa 11	Taiwan	27027340
223	Mak Zai 44-2	Japan		299	Taiwan Okinawa 13	Taiwan	27027342
224	Mak Zai 139-8	Japan	27004834	300	Dayewulong	Taiwan	
225	Mak Zai 138-4	Japan	27004815	301	Tsuyeh	Taiwan	
226	Mak Zai 141-3	Japan	27004857	302	Mak Ct 5320	Taiwan	27004003
227	Mak Zai 2766-2	Japan	27005234	303	Mak Ct 5338	Taiwan	27004006
228	Mak Zai 2763-16	Japan	27005217	304	Mak Ct 5606	Taiwan	

Table 4.1. continued

No.	Accession name	Origin	Accession No.	No.	Accession name	Origin	Accession No.
305	C 3	China	27027355	381	Kna Ck 30	China	27001953
306	C 7	China	27025703	382	Kna Ck 41	China	27001956
307	Kag Cp1	China		383	Kna Ck 49	China	27001958
308	Kag Cp 2	China	27003539	384	Kna Ck 54	China	27001960
309	Mak Cp 1	China	27003513	385	Tieguanyin	China	
310	Mak Cp 2	China	27003514	386	Wuyi	China	
311	Mak Cp 3	China	27003515	387	Kna Cm 3	China	27001967
312	Mak Cp 4	China	27003516	388	Kag Cn 1	China	27003676
313	Mak Cp 8	China	27003520	389	Mak Cn 1	China	27003574
314	Mak Cp 13	China		390	Mak Cn 4	China	27003577
315	Mak Cp 19	China	27003526	391	Mak Cn 6	China	27003579
316	Mak Cp 21	China	27003527	392	Mak Cn 9	China	27003582
317	Mak Cp 24	China	27003530	393	Mak Cn 12	China	27003585
318	Mak Cp 25	China		394	Mak Cn 13	China	27003586
319	Mak Cp 28	China	27003533	395	Mak Cn 14	China	
320	Mak Cp 30	China	27003535	396	Mak Cn 16	China	27003588
321	Kna Cp 13	China	27001973	397	Mak Cn 17	China	27003589
322	Kna Cp 20	China	27001976	398	Mak Cn 21	China	27003592
323	Saikomidori 2	China	27013034	399	Mak Cn 22	China	27003593
324	Karamidori	China		400	Mak Cn 26	China	27003596
325	Mak Cm 1	China	27003540	401	Mak Cn 28	China	27003598
326	Mak Cm 2	China	27003541	402	Mak Cn 29	China	27003599
327	Mak Cm 4	China	27003543	403	Mak Cn 30	China	27003600
328	Mak Cm 5	China	27003544	404	Mak Cn 34	China	27003603
329	Mak Cm 6	China	27003545	405	Mak Cn 35	China	27003604
330	Mak Cm 9	China	27003548	406	Mak Cn 38	China	27003606
331	Mak Cm 12	China	27003551	407	Mak Cn 41	China	27003609
332	Mak Cm 16	China	27003553	408	Mak Cn 42	China	27003610
333	Mak Cm 17	China	27003554	409	Mak Cn 45	China	27003613
334	Mak Cm 19	China	27003555	410	Mak Cn 46	China	27003614
335	Mak Cm 21	China		411	Mak Cn 48	China	27003616
336	Mak Cm 25	China		412	Mak Cn 50	China	27003618
337	Mak Cm 26	China	27003559	413	Mak Cn 51	China	27003619
338	Mak Cm 29	China	27003561	414	Mak Cn 53	China	27003621
339	Mak Cm 31	China	27003563	415	Mak Cn 55	China	27003623
340	Mak Cm 34	China	27003565	416	Mak Cn 56	China	27003624
341	Mak Cm 36	China	27003567	417	Mak Cn 58	China	27003626
342	Mak Cm 37	China	27003568	418	Mak Cn 61	China	27003629
343	Mak Cm 38	China	27003569	419	Mak Cn 64	China	27003632
344	Miya Ck 4	China	27003723	420	Mak Cn 65	China	27003633
345	Miya Ck 10	China		421	Mak Cn 69	China	27003637
346	Miya Ck 19	China	27003725	422	Mak Cn 71	China	27003638
347	Miya Ck 26	China	27025746	423	Mak Cn 73	China	27003640
348	Miya Ck 34	China	27025749	424	Mak Cn 74	China	27003641
349	Miya Ck 35	China	27003726	425	Mak Cn 77	China	27003644
350	Miya Ck 48	China	27003729	426	Mak Cn 78	China	27003645
351	Miya Ck 68	China	27025750	427	Mak Cn 79	China	27003646
352	Miya Ck 70	China		428	Mak Cn 81	China	27003648
353	Miya Ck 80	China	27025752	429	Mak Cn 83	China	27003650
354	Miya Ck 84	China	27003735	430	Mak Cn 84	China	27003651
355	Miya Ck 91	China	27025766	431	Mak Cn 85	China	27003652
356	Miya Ck 95	China	27003738	432	Mak Cn 89	China	27003656
357	Miya Ck 96	China	27003739	433	Mak Cn 91	China	27003658
358	Mak Ck 1	China	27003677	434	Mak Cn 93	China	27003659
359	Mak Ck 2	China	27003678	435	Mak Cn 95	China	27003661
360	Mak Ck 4	China	27003680	436	Mak Cn 98	China	27003663
361	Mak Ck 5	China	27003681	437	Mak Cn 99	China	27003664
362	Mak Ck 6	China	27003682	438	Mak Cn 101	China	27003665
363	Mak Ck 9	China	27003685	439	Mak Cn 105	China	27003668
364	Mak Ck 12	China	27003688	440	Mak Cn 106	China	27003669
365	Mak Ck 13	China	27003689	441	Mak Cn 107	China	27003670
366	Mak Ck 16	China	27003691	442	Mak Cn 110	China	27003672
367	Mak Ck 21	China		443	Kna Cn 1	China	27001941
368	Mak Ck 23	China	27003695	444	Kna Cn 13	China	27001942
369	Mak Ck 25	China		445	Yamanami	China	
370	Mak Ck 28	China	27003699	446	Shizu Cy 1	China	27003760
371	Mak Ck 30	China	27003701	447	Shizu Cy 2	China	27003761
372	Mak Ck 31	China	27003702	448	Shizu Cy 3	China	27003762
373	Mak Ck 32	China	27003703	449	Shizu Cy 4	China	27003763
374	Mak Ck 35	China	27003706	450	Shizu Cy 5	China	27003764
375	Mak Ck 37	China	27003708	451	Shizu Cy 6	China	27003765
376	Mak Ck 39	China	27003710	452	Shizu Cy 8	China	27003766
377	Mak Ck 42	China	27003713	453	Shizu Cy 10	China	27003767
378	Mak Ck 45	China	27003716	454	Shizu Ch 1	China	27003759
379	Mak Ck 46	China	27003717	455	Ch 5340	China	
380	Kna Ck 17	China		456	Mak Ch 5343	China	

Table 4.1. continued

No.	Accession name	Origin	Accession No.	No.	Accession name	Origin	Accession No.
457	Mak Ch 5351	China		533	Mak VtNm 187	Vietnam	
458	Mak Ch 5351	China		534	Mak VtNm 200	Vietnam	
459	Mak Ch 5360	China		535	Mak VtNm 201	Vietnam	
460	Mak Ch 5363	China		536	Mak VtNm 206	Vietnam	
461	Mak Ch 5402	China		537	Thai YG Ehime 1	Thailand	
462	Mak Ch 5410	China		538	Abo 4	Malaysia	27003356
463	Mak Ch 5412	China		539	Abo 18	Malaysia	27003358
464	Karabeni	China	30006309	540	Abo 22	Malaysia	27003360
465	Shizu Ca 1	China	27003768	541	Abo 24	Malaysia	27003362
466	Shizu Ca 2	China	27003769	542	Abo 27	Malaysia	27003364
467	Shizu Ca 3	China	27003770	543	Aj 1	Indonesia	27002782
468	Shizu Ca 4	China	27003771	544	Aj 2	Indonesia	27002783
469	Shizu Ca 5	China	27003772	545	Aj 22	Indonesia	
470	Shizu Ca 6	China	27003773	546	Aj 26	Indonesia	
471	Shizu Ca 7	China	27003774	547	Abo 21	Indonesia	27003359
472	Csi 11	China	27013044	548	Abo 377	Indonesia	27003366
473	Csi 3	China	27013039	549	BUM 1	Myanmar	27003367
474	Csi 7	China	27013041	550	BUM 2	Myanmar	27003368
475	Sichuan 1	China		551	BUM 3	Myanmar	
476	Sichuan 2	China		552	BUM 5	Myanmar	27003371
477	Sichuan 3	China		553	BUM 6	Myanmar	27003372
478	Sichuan 4	China		554	BUM 7	Myanmar	27003373
479	Sichuan 5	China		555	PKS 18	Bangladesh	27003043
480	Sichuan 6	China		556	PKS 43	Bangladesh	27003052
481	Sichuan 7	China		557	PKS 47	Bangladesh	27003053
482	Sichuan 8	China		558	PKS 52	Bangladesh	27003057
483	Sichuan 9	China		559	PKS 75	Bangladesh	27003066
484	Sichuan 10	China		560	PKS 78	Bangladesh	27003068
485	Sichuan 11	China		561	PKS 84	Bangladesh	27003071
486	Yunnan 5	China		562	PKS 89	Bangladesh	27003075
487	Yunnan 6	China		563	PKS 96	Bangladesh	27003079
488	Yunnan 7	China		564	PKS 101	Bangladesh	27003081
489	China 5	China		565	PKS 104	Bangladesh	27003083
490	China 19	China		566	PKS 109	Bangladesh	27003086
491	China TRI 4	China		567	PKS 113	Bangladesh	27003089
492	Shan 35	Vietnam		568	PKS 118	Bangladesh	27003093
493	TR 777	Vietnam		569	PKS 120	Bangladesh	27003095
494	Soizan Daichjaju 3	Vietnam		570	PKS 126	Bangladesh	27003100
495	Sonkau 12	Vietnam		571	PKS 132	Bangladesh	27003102
496	Sonkau 18	Vietnam		572	PKS 135	Bangladesh	27003104
497	Sonkau 30	Vietnam		573	PKS 147	Bangladesh	27003108
498	HazanFactory 11	Vietnam		574	PKS 154	Bangladesh	27003111
499	HazanFactory 14	Vietnam		575	PKS 161	Bangladesh	27003114
500	HazanFactory 20	Vietnam		576	PKS 186	Bangladesh	27003120
501	HazanFactory 3	Vietnam		577	PKS 194	Bangladesh	27003125
502	HazanFactory 9	Vietnam		578	PKS 213	Bangladesh	27003134
503	Vt Dam Dao 3	Vietnam		579	PKS 215	Bangladesh	27003135
504	Ash 1-3	Vietnam	27003376	580	PKS 224	Bangladesh	27003137
505	Shan 21	Vietnam		581	PKS 235	Bangladesh	27003141
506	Shan 37	Vietnam	27003385	582	PKS 236	Bangladesh	
507	Mak VtNm 3	Vietnam		583	PKS 245	Bangladesh	27003145
508	Mak VtNm 10	Vietnam		584	PKS 257	Bangladesh	27003148
509	Mak VtNm 29	Vietnam		585	PKS 267	Bangladesh	27003153
510	Mak VtNm 33	Vietnam		586	PKS 279	Bangladesh	27003157
511	Mak VtNm 48	Vietnam		587	PKS 285	Bangladesh	27003160
512	Mak VtNm 51	Vietnam		588	PKS 286	Bangladesh	27003161
513	Mak VtNm 60	Vietnam		589	PKS 294	Bangladesh	27003167
514	Mak VtNm 70	Vietnam		590	PKS 301	Bangladesh	27003170
515	Mak VtNm 76	Vietnam		591	PKS 317	Bangladesh	27003179
516	Mak VtNm 83	Vietnam		592	PKS 320	Bangladesh	27003181
517	Mak VtNm 93	Vietnam		593	PKS 326	Bangladesh	27003184
518	Mak VtNm 105	Vietnam		594	PKS 334	Bangladesh	27003186
519	Mak VtNm 109	Vietnam		595	PKS 352	Bangladesh	27003192
520	Mak VtNm 116	Vietnam		596	PKS 364	Bangladesh	27003194
521	Mak VtNm 120	Vietnam		597	PKS 379	Bangladesh	27003202
522	Mak VtNm 126	Vietnam		598	PKS 391	Bangladesh	27003206
523	Mak VtNm 136	Vietnam		599	PKS 400	Bangladesh	27003209
524	Mak VtNm 141	Vietnam		600	PKS 402	Bangladesh	
525	Mak VtNm 148	Vietnam		601	PKS 414	Bangladesh	27003214
526	Mak VtNm 153	Vietnam		602	PKS 445	Bangladesh	27003222
527	Mak VtNm 157	Vietnam		603	PKS 446	Bangladesh	27003223
528	Mak VtNm 162	Vietnam		604	PKS 494	Bangladesh	27003239
529	Mak VtNm 166	Vietnam		605	Ai 7	India	27002758
530	Mak VtNm 172	Vietnam		606	Ai 8	India	27002759
531	Mak VtNm 177	Vietnam		607	Ai 16	India	27002762
532	Mak VtNm 182	Vietnam		608	Ai 19	India	

Table 4.1. continued

No.	Accession name	Origin	Accession No.	No.	Accession name	Origin	Accession No.
609	Ai 37	India	27002766	685	Mak Cd 179	India	27003876
610	Ai 86	India	27002771	686	Mak Cd 183	India	27003879
611	Ai 104	India	27002773	687	Mak Cd 184	India	27003880
612	Ai 142	India	27002776	688	Mak Cd 194	India	27003885
613	Ai 173	India	27002781	689	Mak Cd 201	India	27003888
614	Ak 1	India	27003434	690	Mak Cd 206	India	27003891
615	Ak 9	India		691	Mak Cd 207	India	27003892
616	Ak 11	India	27003441	692	Mak Cd 219	India	27003899
617	Ak 14-1	India		693	Mak Cd 221	India	27003900
618	Ak 14-2	India		694	Mak Cd 226	India	27003904
619	Ak 25	India	27003447	695	Mak Cd 239	India	27003910
620	Ak 65	India	27003466	696	Mak Cd 240	India	27003912
621	Ak 94	India	27003478	697	Mak Cd 242	India	27003913
622	Ak 137	India	27003494	698	Mak Cd 247	India	27003916
623	Ak 162	India	27003503	699	Mak Cd 251	India	27003920
624	Ak 216	India	27002792	700	Mak Cd 255	India	27003924
625	Ak 273	India	27002803	701	Mak Cd 257	India	27003926
626	Ak 353	India	27002812	702	Mak Cd 260	India	27003929
627	Ak 433	India	27002829	703	Mak Cd 267	India	27003935
628	Ak 468	India	27002836	704	Mak Cd 269	India	27003937
629	Ak 521	India	27002844	705	Mak Cd 274	India	27003942
630	Ak 532	India	27002848	706	Mak Cd 284	India	27003948
631	Ak 574	India	27002866	707	Mak Cd 289	India	27003952
632	Ak 603	India	27002873	708	Mak Cd 296	India	27003958
633	Ak 649	India	27002876	709	Mak Cd 298	India	27003960
634	Ak 672	India	27002880	710	Mak Cd 303	India	27003965
635	Ak 757	India	27002886	711	Mak Cd 309	India	27003967
636	Ak 844	India	27002890	712	Mak Cd 314	India	27003971
637	Ak 1301	India	27002903	713	Mak Cd 315	India	27003972
638	Ak 1530	India	27002913	714	Mak Cd 316	India	27003973
639	Ak 1612	India	27002920	715	Mak Cd 320	India	27003976
640	Ak 1649	India	27002922	716	Mak Cd 327	India	27003983
641	Ak 1654	India	27002923	717	Mak Cd 333	India	27003988
642	Ak 1699	India	27002929	718	Mak Cd 339	India	27003992
643	Ak 1918	India	27002939	719	Mak Cd 340	India	27003993
644	Ak 2041	India	27002947	720	Mak Cip 3	India	
645	Ak 2306	India	27002960	721	Mak Cip 8	India	
646	Kochi Ak	India	27002963	722	Mak Cip 9	India	
647	Chahon Alu 6	India		723	Mak Cip 11	India	
648	Chahon Ast 37	India		724	Mak Cip 12	India	
649	Kna Cd 49	India	27001987	725	Mak Cip 14	India	
650	Kna Cd 58	India	27001992	726	Mak Cip 19	India	
651	Kna Cd 66	India	27001999	727	Mak Cip 20	India	
652	Kna Cd 137	India	27002020	728	Mak Cip 24	India	
653	Kna Cd 170	India	27002022	729	Mak Cip 28	India	
654	Mak Cd 3	India	27003776	730	Mak Cip 36	India	
655	Mak Cd 5	India	27003778	731	Mak Cip 39	India	
656	Mak Cd 6	India	27003779	732	Mak Cip 43	India	
657	Mak Cd 9	India	27003782	733	Mak Cip 45	India	
658	Mak Cd 14	India	27003783	734	IND 11	India	27002972
659	Mak Cd 25	India		735	IND 18	India	27002976
660	Mak Cd 28	India	27003791	736	IND 24	India	27002982
661	Mak Cd 39	India	27003796	737	IND 39	India	27002988
662	Mak Cd 42	India	27003797	738	IND 40	India	27002989
663	Mak Cd 45	India	27003798	739	IND 49	India	
664	Mak Cd 69	India	27003809	740	IND 67	India	27003000
665	Mak Cd 79	India	27003813	741	IND 75	India	27003005
666	Mak Cd 80	India		742	IND 88	India	27003012
667	Mak Cd 86	India		743	IND 103	India	27003022
668	Mak Cd 88	India	27003818	744	IND 112	India	27003027
669	Mak Cd 102	India	27003825	745	IND 113	India	27003028
670	Mak Cd 105	India	27003828	746	IND 117	India	27003031
671	Mak Cd 114	India	27003832	747	SMP 20	India	27003351
672	Mak Cd 120	India	27003834	748	Stock 808	India	27003347
673	Mak Cd 129	India	27003841	749	Kangura 70	India	27002068
674	Mak Cd 131	India	27003843	750	Aindi 2	India	
675	Mak Cd 133	India	27003845	751	Aindi 4	India	
676	Mak Cd 138	India	27003848	752	A 4	India	27002042
677	Mak Cd 141	India	27003851	753	Chahon Ace 37	Sri Lanka	
678	Mak Cd 142	India	27003852	754	SL 1-2	Sri Lanka	
679	Mak Cd 149	India	27003857	755	SRL 3	Sri Lanka	27003391
680	Mak Cd 156	India	27003862	756	SL 12-6	Sri Lanka	
681	Mak Cd 157	India	27003863	757	SL 15-12	Sri Lanka	
682	Mak Cd 167	India	27003869	758	SL 15-19	Sri Lanka	
683	Mak Cd 168	India	27003870	759	SL 15-28	Sri Lanka	
684	Mak Cd 174	India	27003873	760	SL 15-4	Sri Lanka	

Table 4.1. continued

No.	Accession name	Origin	Accession No.
761	SL 15-7	Sri Lanka	
762	SL 15-50	Sri Lanka	
763	SL 15-60	Sri Lanka	
764	SL 5-11	Sri Lanka	
765	SL 5-21	Sri Lanka	
766	SL 5-31	Sri Lanka	
767	SL 6-2	Sri Lanka	
768	SRL 10	Sri Lanka	27003397
769	SRL 16	Sri Lanka	27003402
770	SRL 30	Sri Lanka	27003407
771	SRL 41	Sri Lanka	27003413
772	SRL 76	Sri Lanka	27003422
773	IRN 1	Iran	
774	IRN 2	Iran	
775	IRN 7	Iran	
776	IRN 10	Iran	
777	IRN 14	Iran	
778	IRN 23	Iran	
779	IRN29	Iran	
780	IRN 38	Iran	
781	CR 22	Georgia	27002075
782	CR 24	Georgia	27002077
783	CR 28	Georgia	27002081
784	MC 1	Georgia	
785	MC 17	Georgia	
786	MC 28	Georgia	
787	MC 30	Georgia	
788	MC 43	Georgia	
761	SL 15-7	Sri Lanka	
762	SL 15-50	Sri Lanka	
763	SL 15-60	Sri Lanka	
764	SL 5-11	Sri Lanka	
765	SL 5-21	Sri Lanka	
766	SL 5-31	Sri Lanka	
767	SL 6-2	Sri Lanka	
768	SRL 10	Sri Lanka	27003397
769	SRL 16	Sri Lanka	27003402
770	SRL 30	Sri Lanka	27003407
771	SRL 41	Sri Lanka	27003413
772	SRL 76	Sri Lanka	27003422
773	IRN 1	Iran	
774	IRN 2	Iran	
775	IRN 7	Iran	
776	IRN 10	Iran	
777	IRN 14	Iran	
778	IRN 23	Iran	
779	IRN29	Iran	
780	IRN 38	Iran	
781	CR 22	Georgia	27002075
782	CR 24	Georgia	27002077
783	CR 28	Georgia	27002081
784	MC 1	Georgia	
785	MC 17	Georgia	
786	MC 28	Georgia	
787	MC 30	Georgia	
788	MC 43	Georgia	

Table 4.2. Primer sequences and accession numbers of SSR markers used in this study

Marker name	Accssion number	Forward primer sequence	Reverse primer sequence
MSG0258	AB624013	actcatcaccatgccttctccatc	gtttagctcaactggtggaacctcaact
MSG0361	AB624038	agatggaggtagagagagaggcag	gtttgtccctctcattttcaacgc
MSE0173	AB623905	gtgttcaccaacaactcaccaagg	tgtcgaacaaagatacaccccaaa
MSG0429	AB624066	aggaccgttcttccctacctgtaa	gtttgagattgaggatgtggtcgttgt
MSE0029	FS949897	atagccaatcaagctcctcctcct	agtctgttcctcccttgatgatcg
MSG0533	AB624117	agacctagccaagacaaccacacc	gtttcccctattttcccgactgtct
MSE0250	AB623937	cttccccaaaccaccatcaaaata	gaaattgaagaacacgaacctgcc
MSG0380	AB624040	acagaccttcaccctctccatttc	gtttacctctgccttcgttcttcagc
MSE0313	AB623968	tgctatgccgcctaacaaaaactt	accaccaacaacaattcccactct
MSG0702	AB624201	atgttggaccagtagcaccgaaac	gtttcgggttcctttctcaaactcc
MSG0423	AB624063	actccatgtgctgctctgtagttc	gtttgcaggaagttgagccagac
MSG0610	AB624154	acagaggaggaagatgatcggtaa	gtttgaagaagaagaaaactcccgccat
MSE0108	FS948805	agtccatggtggttgatgatccttt	ttgggagtaggattcttgcagagc
MSG0403	AB624054	atgatcgccggtttagagatgaat	gtttaagctggctaacctacacggagc
MSE0237	AB623931	ctctccttcttcacaccctccaaa	ttgttctcaaagaacctccttcgc
MSE0291	AB623953	aatcaaataacacttgcacccgc	aaaaagagaaagtcacgtccacgg
MSE0113	AB485972	taccttctgcaactccagcaatcc	tgagattgaccatctttcatcgga
MSE0143	FS951913	gccttttgtcagaaacggtgact	cagcaatctttggttttgttgtgc
MSG0681	AB624190	agggtttgcgtcttcaaagagaga	gtttgtaacacttgccacgtttcg
MSE0107	AB485971	tctctctactcctgcgcaatctca	tcaaagatgttgctctcgtcaacc
MSE0083	AB623985	gaggaaagagattatgcggtgtgg	gtgagccttcaaaagacagcaacg
MSG0470	AB624089	atagggttcgaaaatggcagg	gtttgaggtggcaagtttgtgactgt
MSG0699	AB624199	atgcgacagtgttgctgagatttt	gtttcaaaaatggggtgtctacagaggg

Table 4.3. Genetic diversity indices for 23 SSR marker loci

Marker name	LG[a]	repeat region	size range	Allele No	*He*[b]	*Ho*[c]	PIC[d]
MSG0258	LG01	(tc)25(ta)4,(ta)3	238 - 286	28	0.90	0.82	0.89
MSG0361	LG01	(ag)26	239 - 293	40	0.90	0.54	0.90
MSE0173	LG02	(tc)10(ta)9	277 - 322	37	0.90	0.74	0.90
MSG0429	LG02	(tc)24	258 - 317	27	0.83	0.63	0.82
MSE0029	LG03	(ag)14,(ag)7,(gga)3	357 - 398	25	0.89	0.66	0.88
MSG0533	LG03	(ag)18,(ga)3	192 - 240	35	0.91	0.63	0.90
MSE0250	LG04	(atc)3,(ca)3,(tc)19	320 - 363	27	0.87	0.75	0.86
MSG0380	LG04	(ag)21,(ta)4	245 - 281	21	0.88	0.71	0.87
MSE0313	LG05	(ag)3,(ag)12	215 - 259	30	0.85	0.67	0.84
MSG0702	LG05	(ag)18	250 - 295	27	0.87	0.63	0.86
MSG0423	LG06	(tc)4,(tc)13(ac)7	141 - 178	23	0.88	0.48	0.87
MSG0610	LG06	(tc)16,(tg)3	270 - 302	15	0.82	0.64	0.80
MSE0108	LG07	(tc)6(ta)8	245 - 277	20	0.81	0.54	0.78
MSG0403	LG07	(tc)25	250 - 300	24	0.87	0.65	0.86
MSE0237	LG08	(acc)3,(tc)4,(tc)3,(tc)4,(tct)5	331 - 362	14	0.57	0.41	0.55
MSE0291	LG08	(ca)3,(tc)9	230 - 266	26	0.84	0.68	0.82
MSE0113	LG09	(tc)14	333 - 396	34	0.91	0.73	0.90
MSE0143	LG09	(ag)8,(tc)4	303 - 345	28	0.83	0.73	0.82
MSG0681	LG10	(ag)18	205 - 248	21	0.89	0.69	0.88
MSE0107	LG11	(tc)8,(ca)3,(cct)3	279 - 321	24	0.88	0.72	0.87
MSE0083	LG12	(tct)4(ct)3,(tc)3,(tct)6	228 - 267	27	0.85	0.74	0.83
MSG0470	LG13	(ga)3,(ag)20	144 - 185	24	0.90	0.47	0.89
MSG0699	LG15	(ag)18	236 - 292	42	0.90	0.60	0.89
Mean	-	-	-	26.9	0.86	0.62	0.85

[a] LG, Linkage group .

[b] *He*, expected heterozygosity.

[c] *Ho*, observed heterozygosity.

[d] PIC, polymorphism information content.

Table 4.4. Genetic diversity statistics of tea germplasms.

Origin / core collection	Sample size (no. of accessions in the whole NIVTS collection)	No. of alleles	H_E[a]	H_O[b]	PIC[c]	Allelic richness[d]
Japan	238 (5065)	340	0.70	0.55	0.67	3.29
Korea	30 (731)	205	0.77	0.62	0.74	3.79
Taiwan	36 (167)	317	0.86	0.62	0.84	4.49
China	187 (390)	436	0.85	0.69	0.84	4.36
Vietnam	45 (271)	299	0.78	0.54	0.76	3.99
Thailand	1 (1)	36	ND[e]	ND[e]	0.39	ND[e]
Malaysia	5 (10)	111	0.70	0.61	0.66	3.88
Indonesia	6 (10)	110	0.69	0.61	0.64	3.61
Myanmar	6 (7)	111	0.66	0.54	0.62	3.51
Bangladesh	50 (202)	303	0.79	0.66	0.77	4.06
India	148 (679)	395	0.85	0.65	0.84	4.37
Sri Lanka	20 (169)	244	0.82	0.66	0.80	4.28
Iran	8 (33)	164	0.79	0.78	0.76	4.19
Georgia	8 (37)	155	0.79	0.68	0.76	4.22
Core 24	24	368 (59.5%)[f]	0.89	0.74	0.88	4.88
Core 48	48	462 (74.6%)[f]	0.89	0.70	0.88	4.82
Core 96	96	546 (88.2%)[f]	0.89	0.66	0.88	4.73
Core 192	192	616 (99.5%)[f]	0.88	0.64	0.87	4.64
All accessions	788 (7781)	619	0.86	0.62	0.85	4.41

[a] H_E, expected heterozygosity.

[b] H_O, observed heterozygosity.

[c] PIC, polymorphism information content.

[d] Allelic richness based on a minimum sample size of three diploid individuals.

[e] Not determined.

[f] Numbers shown in parentheses are coverage of allele numbers by each core collection.

Table 4.5. Summary of AMOVA based on 23 SSR loci.

Source	df	Sum of squares	Mean squares	Variance components	% of total variation	Statistic		*P*-value
Among populations	12	2520.482	210.040	3.553	14			
Within populations	774	17395.617	22.475	22.475	86	PhiPT [a]	0.137	0.001
Total	786	19916.099		26.028	100			

[a] PhiPT、proportion of the total genetic variance among populations.

Table 4.6. Pairwise *Fst* values between populations

	Japan	Korea	Taiwan	China	Vietnam	Malaysia	Indonesia	Myanmar	Bangladesh	India	Sri Lanka	Iran	Georgia
Japan		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Korea	0.116		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Taiwan	0.135	0.081		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
China	0.091	0.033	0.048		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Vietnam	0.190	0.141	0.104	0.097		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Malaysia	0.162	0.106	0.071	0.068	0.091		0.027	0.002	0.041	0.003	0.054	0.001	0.001
Indonesia	0.182	0.146	0.092	0.086	0.107	0.038		0.001	0.002	0.001	0.001	0.001	0.001
Myanmar	0.217	0.163	0.103	0.114	0.127	0.076	0.073		0.001	0.001	0.001	0.001	0.001
Bangladesh	0.166	0.132	0.090	0.085	0.099	0.011	0.021	0.070		0.001	0.004	0.001	0.001
India	0.107	0.054	0.044	0.022	0.083	0.023	0.045	0.083	0.039		0.001	0.001	0.001
Sri Lanka	0.140	0.091	0.058	0.051	0.085	0.014	0.026	0.071	0.008	0.013		0.001	0.001
Iran	0.132	0.062	0.047	0.025	0.110	0.068	0.113	0.134	0.089	0.018	0.043		0.005
Georgia	0.137	0.068	0.052	0.027	0.102	0.070	0.098	0.130	0.098	0.029	0.063	0.020	

Fst Values below diagonal. Probability which random values are greater than or equal to the observed value based on 999 permutations is shown above diagonal.

Table 4.7. List of core collections of tea germplasms

No.	Accession name	ID[a]	Type	Country or region	Origin detail	Core24	Core48	Core96	Core192
1	Yabukita	168695	cultivar	Japan	Shizuoka	√	√	√	√
2	Sayamakaori	168733	cultivar	Japan	Saitama	√	√	√	√
3	Minamisayaka	178580	cultivar	Japan	Miyazaki	√	√	√	√
4	Shizu Inzatsu 131	151825	clonal line	Japan	Shizuoka	√	√	√	√
5	Sayamamidori	168730	cultivar	Japan	Shizuoka				√
6	Z 1	178537	clonal line	Japan	NIVTS				√
7	Shizu Zai 16	178561	clonal line	Japan	Shizuoka				√
8	Kin Zai Aki 91-17	250941	landrace	Japan	Akita			√	√
9	Kin Zai Aki 91-28	250952	landrace	Japan	Akita				√
10	Mak Zai 92-9	122007	landrace	Japan	Fukui				√
11	Gifu Zai 30-8	250622	landrace	Japan	Gifu			√	√
12	NG 10	120159	landrace	Japan	Shizuoka				√
13	Shizu At 10	251005	landrace	Japan	Shizuoka				√
14	Shizu At 14	251009	landrace	Japan	Shizuoka		√	√	√
15	Mak Zai 69-13	121994	landrace	Japan	Mie				√
16	Mak Zai 66-2	121963	landrace	Japan	Mie				√
17	Mak Zai 70-4	121995	landrace	Japan	Mie				√
18	Mie At 120	120465	landrace	Japan	Mie		√	√	√
19	Kna Zai Shig 93	120212	landrace	Japan	Shiga				√
20	Kna Zai Kyo 88-135	131926	landrace	Japan	Kyoto				√
21	Mak Zai 147-30	122899	landrace	Japan	Kyoto			√	√
22	Mak Zai 113-3	122295	landrace	Japan	Hyogo				√
23	Mak Zai 27-1	122302	landrace	Japan	Hiroshima				√
24	Fukuoka Zai 26-15	250624	landrace	Japan	Fukuoka				√
25	Mak Zai 39-18	122452	landrace	Japan	Fukuoka				√
26	Mak Zai 120-4	151924	landrace	Japan	Saga			√	√
27	Kna Zai Tsushima 19-1	239439	landrace	Japan	Nagasaki				√
28	Kna Zai Goto 7-2	239452	landrace	Japan	Nagasaki			√	√
29	Kna Zai Goto 2-4	239449	landrace	Japan	Nagasaki		√	√	√
30	Kna Zai Tsushima 1-7	239379	landrace	Japan	Nagasaki		√	√	√
31	Kna Zai Tsushima 2-4	239383	landrace	Japan	Nagasaki				√
32	Itsuki Yamacha Kozuru 32-14	250625	wild tea	Japan	Kumamoto				√
33	Mak Zai 2824-2	123040	wild tea	Japan	Kumamoto				√
34	Mak Zai 2948-3	123071	landrace	Japan	Oita				√
35	ME 52	137548	clonal line	Japan	Miyazaki				√
36	Mak Zai 138-4	122580	landrace	Japan	Miyazaki			√	√
37	Mak Zai 2763-16	122982	wild tea	Japan	Miyazaki				√
38	Kna Korea 93-1-48	139532	landrace	Korea	South Jeolla				√
39	Kna Korea 93-4-33	139590	landrace	Korea	South Jeolla				√
40	Kna Korea 93-5-33	139620	landrace	Korea	South Jeolla			√	√
41	Kna Korea 93-23-4	140080	landrace	Korea	South Jeolla			√	√
42	Kna Korea 93-29-14	140183	landrace	Korea	South Jeolla				√
43	Mak Cok 113	250626	landrace	Taiwan	Okinawa			√	√
44	Mak Cok 129	250627	landrace	Taiwan	Okinawa	√	√	√	√
45	Ay 5315	121109	wild tea	Taiwan	Nantou				√
46	Taiwan Yamacha 25	121037	wild tea	Taiwan	Kaohsiung	√	√	√	√
47	Taiwan Yamacha 37	121048	wild tea	Taiwan	Kaohsiung	√	√	√	√
48	Taiwan Yamacha 40	121051	wild tea	Taiwan	Kaohsiung		√	√	√
49	Taiwan Yamacha 47	121058	wild tea	Taiwan	Kaohsiung				√
50	Taiwan Yamacha 53	121064	wild tea	Taiwan	Kaohsiung		√	√	√
51	Taiwan Yamacha 67	121076	wild tea	Taiwan	Kaohsiung				√
52	Taiwan Yamacha 73	121082	wild tea	Taiwan	Kaohsiung	√	√	√	√
53	Taiwan Yamacha 84	121093	wild tea	Taiwan	Kaohsiung				√
54	Taiwan Yamacha 88	121097	wild tea	Taiwan	Kaohsiung			√	√
55	Kna Ay 94	128723	landrace	Taiwan	Nantou				√
56	Taiwan Yamacha 18	121031	wild tea	Taiwan	TES[b]			√	√
57	Taiwan Yamacha Nishi1	250628	wild tea	Taiwan	TES[b]			√	√
58	Taiwan Yamacha 95	121104	wild tea	Taiwan	TES[b]			√	√
59	Tsuyeh	120083	cultivar	Taiwan	Taipei				√
60	Mak Ct 5338	121774	landrace	Taiwan	Unknown	√	√	√	√
61	Mak Ct 5606	121784	landrace	Taiwan	Unknown	√	√	√	√
62	C 7	137527	landrace	China	Zhejiang				√
63	Kag Cp 2	121307	landrace	China	Zhejiang				√
64	Mak Cp 1	121281	landrace	China	Zhejiang			√	√
65	Mak Cp 2	121282	landrace	China	Zhejiang	√	√	√	√
66	Mak Cp 4	121284	landrace	China	Zhejiang		√	√	√
67	Mak Cp 25	151887	landrace	China	Zhejiang				√
68	Kna Cp 13	169206	landrace	China	Jiangxi				√
69	Saikomidori 2	128627	landrace	China	Zhejiang			√	√
70	Mak Cm 2	121309	landrace	China	Anhui				√
71	Mak Cm 5	121312	landrace	China	Anhui				√
72	Mak Cm 19	121323	landrace	China	Anhui			√	√
73	Mak Cm 25	151880	landrace	China	Anhui				√
74	Mak Cm 36	121335	landrace	China	Anhui				√
75	Mak Cm 38	121337	landrace	China	Anhui				√
76	Miya Ck 26	137570	landrace	China	Anhui	√	√	√	√

Table 4.7. continued

No.	Accession name	ID[a]	Type	Country or region	Origin detail	Core24	Core48	Core96	Core192
77	Miya Ck 84	121503	landrace	China	Anhui			√	√
78	Miya Ck 91	137590	landrace	China	Anhui			√	√
79	Mak Ck 2	121446	landrace	China	Anhui		√	√	√
80	Mak Ck 6	121450	landrace	China	Anhui				√
81	Mak Ck 13	121457	landrace	China	Anhui			√	√
82	Mak Ck 28	121467	landrace	China	Anhui			√	√
83	Mak Ck 32	121471	landrace	China	Anhui	√	√	√	√
84	Mak Ck 37	121476	landrace	China	Anhui		√	√	√
85	Kna Ck 17	169210	landrace	China	Anhui	√	√	√	√
86	Kna Ck 30	169187	landrace	China	Anhui				√
87	Kna Ck 54	169194	landrace	China	Anhui			√	√
88	Kag Cn 1	121444	landrace	China	Jiangxi				√
89	Mak Cn 6	121347	landrace	China	Jiangxi				√
90	Mak Cn 9	121350	landrace	China	Jiangxi			√	√
91	Mak Cn 21	121360	landrace	China	Jiangxi			√	√
92	Mak Cn 29	121367	landrace	China	Jiangxi				√
93	Mak Cn 30	121368	landrace	China	Jiangxi		√	√	√
94	Mak Cn 35	121372	landrace	China	Jiangxi				√
95	Mak Cn 64	121400	landrace	China	Jiangxi				√
96	Mak Cn 78	121413	landrace	China	Jiangxi		√	√	√
97	Mak Cn 85	121420	landrace	China	Jiangxi				√
98	Kna Cn 13	169178	landrace	China	Jiangxi		√	√	√
99	Shizu Cy 2	121529	landrace	China	Hubei				√
100	Shizu Cy 4	121531	landrace	China	Hubei				√
101	Shizu Cy 5	121532	landrace	China	Hubei		√	√	√
102	Shizu Cy 6	121533	landrace	China	Hubei				√
103	Shizu Cy 8	121534	landrace	China	Hubei		√	√	√
104	Shizu Cy 10	121535	landrace	China	Hubei				√
105	Mak Ch 5343	121511	landrace	China	Hubei			√	√
106	Mak Ch 5351	121514	landrace	China	Hubei				√
107	Mak Ch 5360	121518	landrace	China	Hubei				√
108	Mak Ch 5363	121520	landrace	China	Hubei			√	√
109	Mak Ch 5402	121523	landrace	China	Hubei				√
110	Mak Ch 5412	121526	landrace	China	Hubei		√	√	√
111	Karabeni	168750	cultivar	China	Hubei		√	√	√
112	Csi 7	128634	landrace	China	Sichuan			√	√
113	Sichuan 1	250629	landrace	China	Sichuan				√
114	Sichuan 2	250630	landrace	China	Sichuan			√	√
115	Sichuan 6	250631	landrace	China	Sichuan				√
116	Sichuan 7	250632	landrace	China	Sichuan			√	√
117	Sichuan 8	250633	landrace	China	Sichuan			√	√
118	Sichuan 10	250634	landrace	China	Sichuan	√	√	√	√
119	Yunnan 7	250635	landrace	China	Yunnan				√
120	Sonkau 18	250636	landrace	Vietnam	Yen Bai				√
121	HazanFactory 11	250637	landrace	Vietnam	Ha Giang			√	√
122	HazanFactory 14	250638	landrace	Vietnam	Ha Giang				√
123	HazanFactory 20	250639	landrace	Vietnam	Ha Giang	√	√	√	√
124	Vt Dam Dao 3	250640	landrace	Vietnam	Vinh Phuc				√
125	Ash 1-3	121144	landrace	Vietnam	Unknown			√	√
126	Shan 21	121149	landrace	Vietnam	Unknown	√	√	√	√
127	Mak VtNm 10	250641	landrace	Vietnam	Tuyen Quan				√
128	Mak VtNm 33	250642	landrace	Vietnam	Tuyen Quan				√
129	Mak VtNm 60	250643	landrace	Vietnam	Tuyen Quan				√
130	Mak VtNm 70	250644	landrace	Vietnam	Yen Bai				√
131	Mak VtNm 83	250645	landrace	Vietnam	Yen Bai			√	√
132	Mak VtNm 109	250646	landrace	Vietnam	Thai Nguyen			√	√
133	Mak VtNm 120	250647	landrace	Vietnam	Thai Nguyen		√	√	√
134	Mak VtNm 126	250648	landrace	Vietnam	Thai Nguyen			√	√
135	Mak VtNm 136	250649	landrace	Vietnam	Thai Nguyen	√	√	√	√
136	Mak VtNm 153	250650	landrace	Vietnam	Phu Tho				√
137	Mak VtNm 157	250651	landrace	Vietnam	Phu Tho				√
138	Mak VtNm 172	250652	landrace	Vietnam	Phu Tho		√	√	√
139	Mak VtNm 182	250653	landrace	Vietnam	Phu Tho				√
140	Mak VtNm 201	250654	landrace	Vietnam	Thai Nguyen				√
141	Mak VtNm 206	250655	landrace	Vietnam	Phu Tho			√	√
142	Thailand 1	250656	landrace	Thailand	Unknown	√	√	√	√
143	Abo 4	121124	landrace	Malaysia	Kuala Lumpur			√	√
144	Abo 27	121132	landrace	Malaysia	Kuala Lumpur			√	√
145	Aj 2	120552	landrace	Indonesia	Unknown				√
146	Abo 21	121127	landrace	Indonesia	Kuala Lumpur			√	√
147	BUM 5	121139	landrace	Myanmar	Shan		√	√	√
148	BUM 7	121141	landrace	Myanmar	Shan				√
149	PKS 104	120852	landrace	Bangladesh	Sylhet				√
150	PKS 113	120858	landrace	Bangladesh	Sylhet				√
151	PKS 126	120869	landrace	Bangladesh	Sylhet			√	√
152	PKS 135	120873	landrace	Bangladesh	Sylhet	√	√	√	√
153	PKS 147	120877	landrace	Bangladesh	Sylhet				√

Table 4.7. continued

No.	Accession name	ID[a]	Type	Country or region	Origin detail	Core24	Core48	Core96	Core192
154	PKS 161	120883	landrace	Bangladesh	Sylhet			√	√
155	PKS 186	120889	landrace	Bangladesh	Sylhet				√
156	PKS 215	120904	landrace	Bangladesh	Sylhet				√
157	PKS 224	120906	landrace	Bangladesh	Sylhet			√	√
158	Ak 14-1	250657	landrace	India	Darjeeling				√
159	Ak 273	120572	landrace	India	Darjeeling				√
160	Ak 353	120581	landrace	India	Darjeeling		√	√	√
161	Ak 603	120642	landrace	India	Darjeeling			√	√
162	Ak 757	120655	landrace	India	Darjeeling				√
163	Ak 1301	120672	landrace	India	Darjeeling			√	√
164	Ak 1649	120691	landrace	India	Darjeeling				√
165	Kna Cd 58	169225	landrace	India	Darjeeling				√
166	Kna Cd 66	169232	landrace	India	Darjeeling				√
167	Kna Cd 137	172693	landrace	India	Darjeeling				√
168	Mak Cd 14	121551	landrace	India	Darjeeling		√	√	√
169	Mak Cd 45	121566	landrace	India	Darjeeling				√
170	Maku Cd 80	151892	landrace	India	Darjeeling				√
171	Mak Cd 120	121602	landrace	India	Darjeeling				√
172	Mak Cd 138	121616	landrace	India	Darjeeling				√
173	Mak Cd 221	121668	landrace	India	Darjeeling	√	√	√	√
174	Mak Cd 267	121703	landrace	India	Darjeeling		√	√	√
175	Mak Cd 298	121728	landrace	India	Darjeeling				√
176	Mak Cd 320	121744	landrace	India	Darjeeling	√	√	√	√
177	Mak Cip 28	250659	landrace	India	Gujarat				√
178	Mak Cip 43	250660	landrace	India	Gujarat				√
179	Aindi 4	250661	landrace	India	Assam			√	√
180	SL 12-6	250662	landrace	Sri Lanka	Ratnapura		√	√	√
181	SL 15-12	250663	landrace	Sri Lanka	Kyandy				√
182	SL 15-4	250664	landrace	Sri Lanka	Kyandy		√	√	√
183	SL 15-50	250665	landrace	Sri Lanka	Kyandy				√
184	SL 15-60	250666	landrace	Sri Lanka	Kyandy	√	√	√	√
185	SL 6-2	250667	landrace	Sri Lanka	Badulla			√	√
186	SRL 30	121175	landrace	Sri Lanka	Unknown			√	√
187	SRL 41	121181	landrace	Sri Lanka	Unknown				√
188	IRN 2	121820	landrace	Iran	Gilan			√	√
189	IRN 10	121828	landrace	Iran	Gilan			√	√
190	IRN 23	121836	landrace	Iran	Gilan				√
191	MC 17	121857	landrace	Georgia	Unknown				√
192	MC 30	121862	landrace	Georgia	Unknown	√	√	√	√

a ID of the Gene Bank, National Institute of Agrobiological Science, Japan
a Tea Research and Extension Station

Table 4.8. Comparison of CVs and variances (σ^2) of chemical constituents of young shoots among core collections.

		All accessions	Core 192	Core 96	Core 48	Core 24	Japan
Total nitrogen	CV	9.7	10.1	10.4	9.1	7.8	10.6
	σ^2	0.287	0.315	0.343	0.273	0.198	0.315
Free amino acids	CV	31.3	33.9	35.9	31.5	33.7	26.7
	σ^2	0.701	0.786	0.906	0.818	0.873	0.569 *
Theanine	CV	38.8	43.9	47.6	42.3	44.6	33.8
	σ^2	0.233	0.285 *	0.365 *	0.356 *	0.353 *	0.200
Neutral detergent fiber	CV	13.6	13.3	13.2	11.8	10.1	11.6
	σ^2	5.220	4.890	4.750	3.760	2.840	4.440
Tannin	CV	13.0	13.5	14.6	14.2	15.1	10.8
	σ^2	4.710	5.240	6.040	5.410	6.270	2.970
Caffeine	CV	20.0	18.7	17.6	15.7	14.9	14.4
	σ^2	0.430	0.400	0.373	0.297	0.273	0.156
Vitamin C	CV	21.9	20.5	19.6	17.9	17.0	12.5
	σ^2	0.00977	0.00835	0.00735	0.00613	0.00538	0.00384 *

* Significant in the *F*-test for equality of variances at $p < 0.05$.

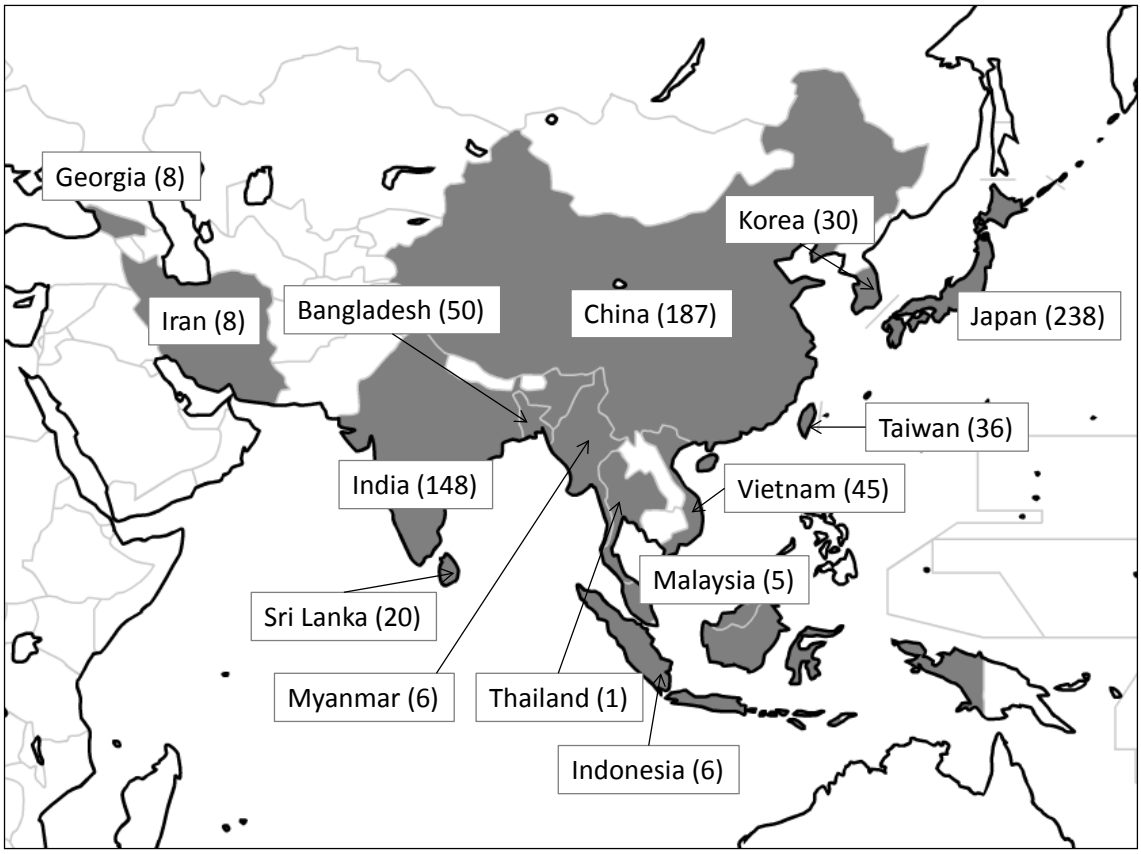

Fig. 4.1 Origin of germplasms analyzed in this study. The numbers of accessions analyzed are shown in parentheses.

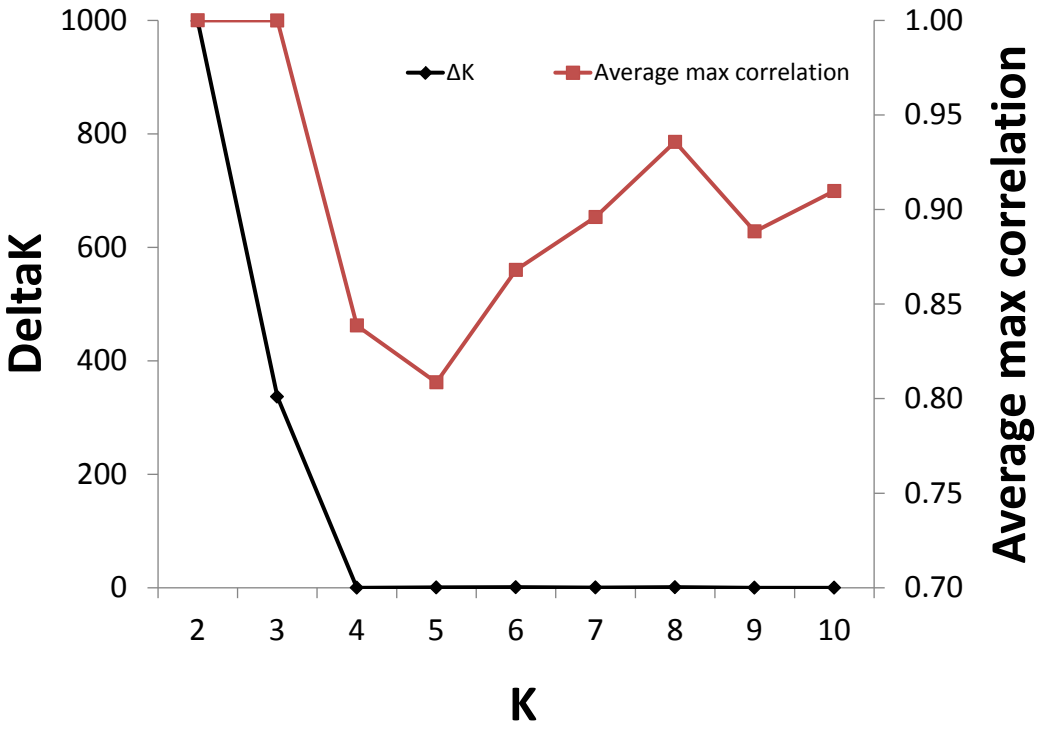

Fig 4.2. DeltaK values and average maximum correlations in the Structure result

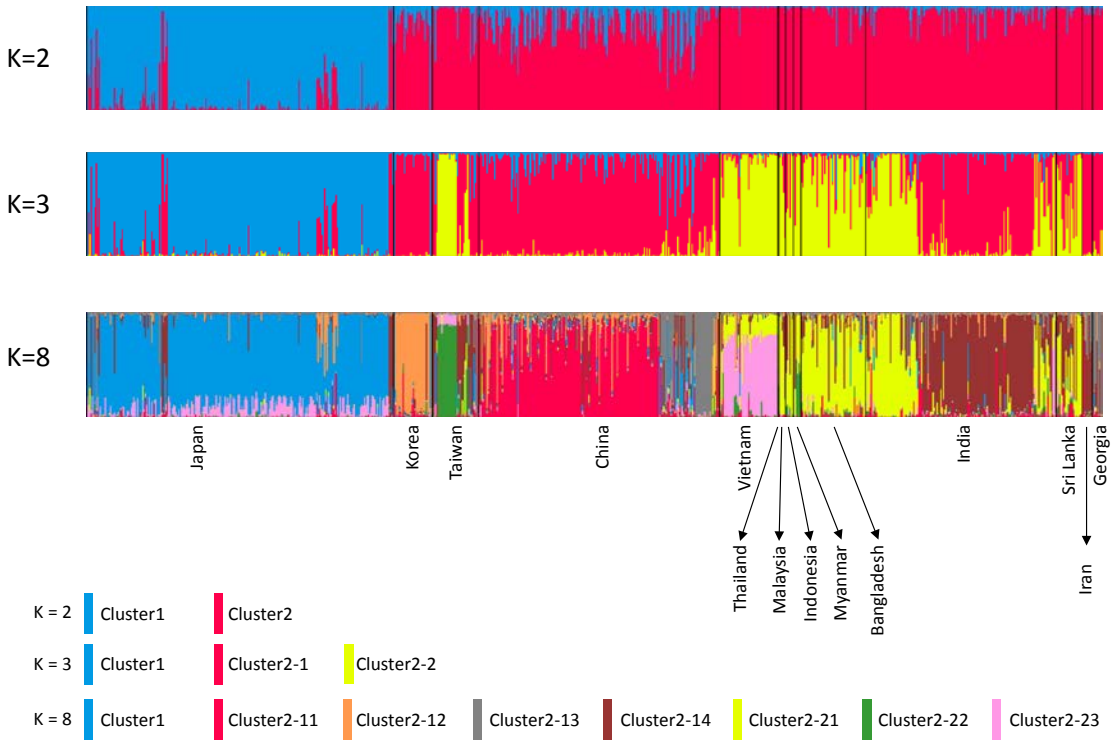

Fig. 4.3. Structure analysis of 788 tea germplasm accessions analyzed with 23 SSR marker loci. Colors represent the inferred ancestral populations. Results for *K* = 2, 3, and 8 are shown.

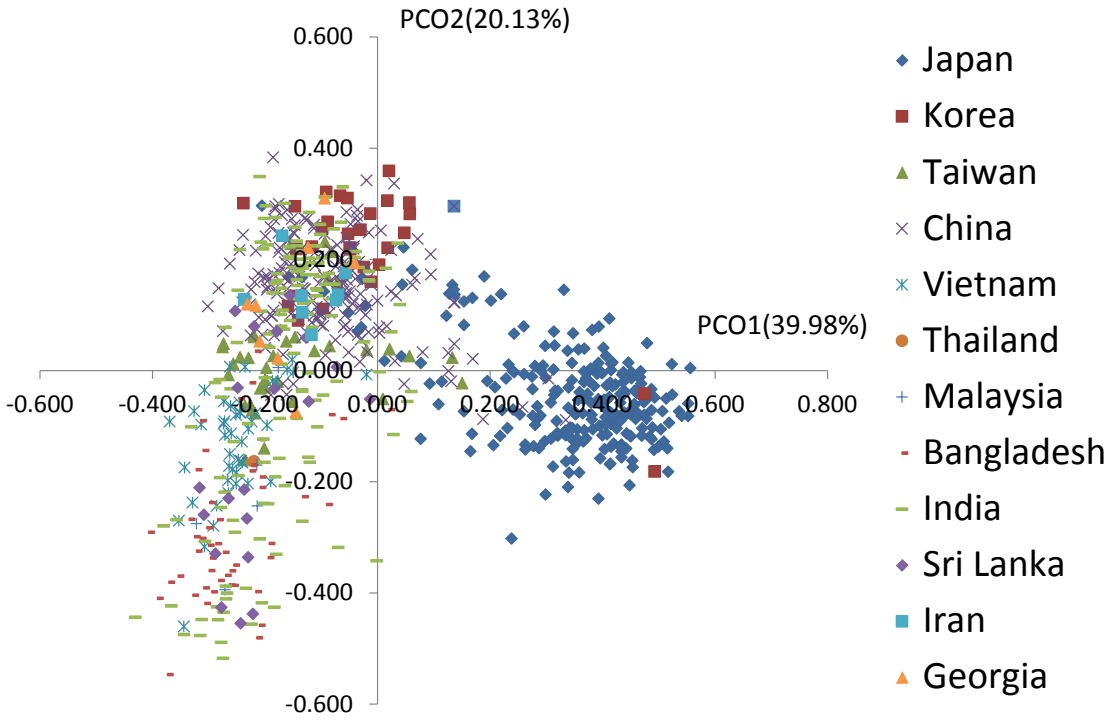

Fig. 4.4. Principal-coordinates analysis (PCO) plot using Nei's genetic distance matrix (D_s: Nei 1978). The first two axes together account for 60% of the overall variability.

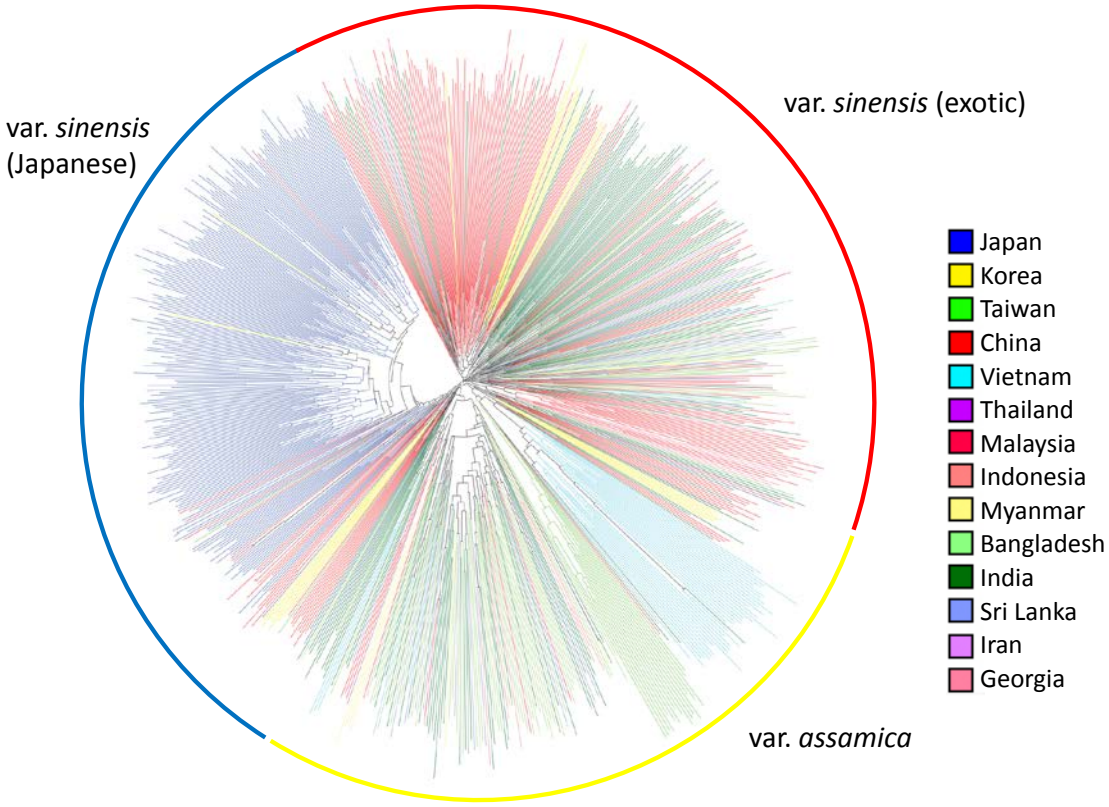

Fig. 4.5. Neighbor-joining tree for the 788 tea germplasm accessions based on data for 23 SSR markers.

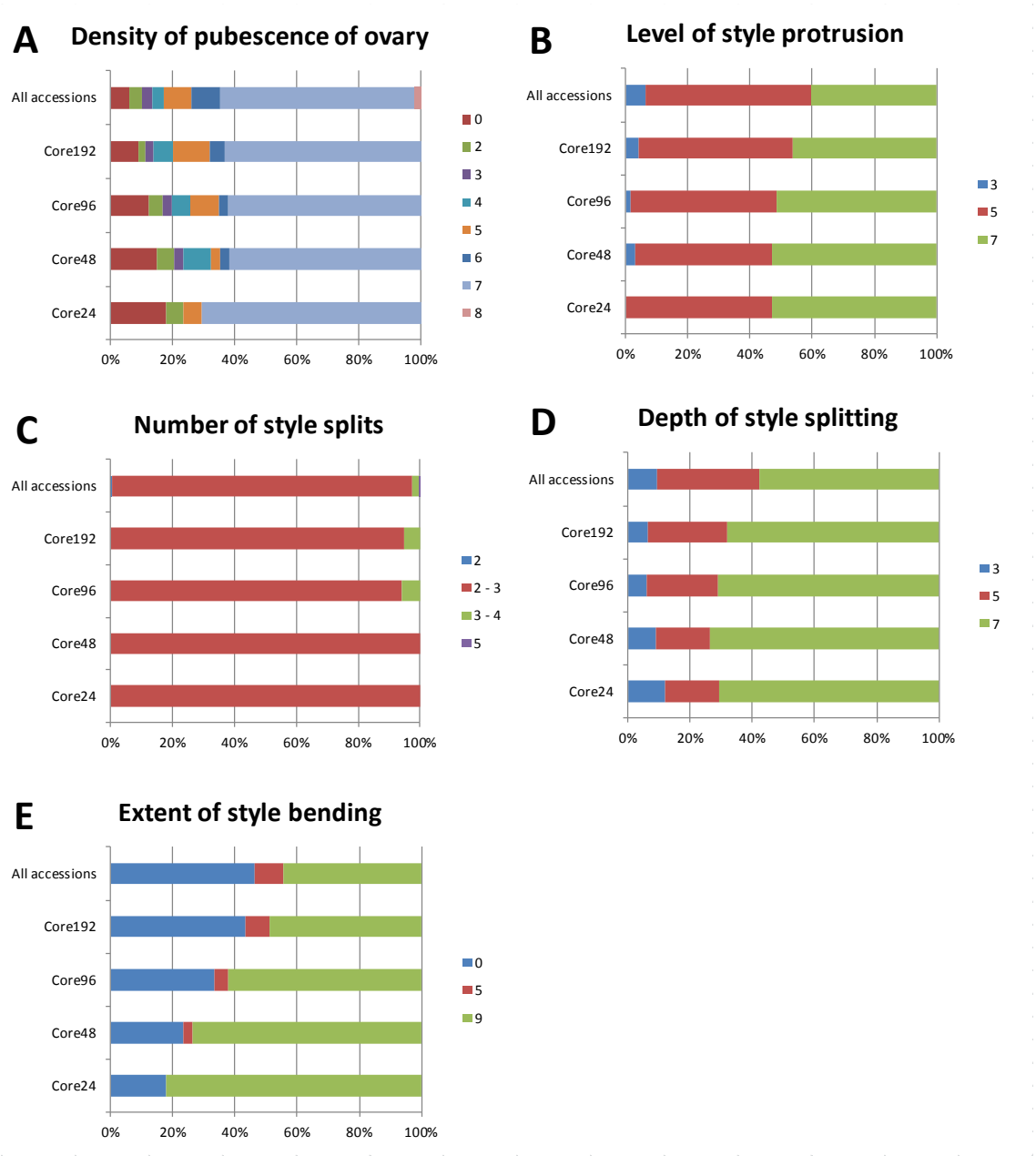

Fig. 4.6. Distribution of floral morphologies of the tea germplasms and the core collections. **A**. Density of pubescence of ovary: 0, no pubescence; 2, very little; 3, little; 4, little to moderate; 5, moderate; 6, slightly dense; 7, dense; 8, very dense. **B**. Level of style protrusion: 3, style < stamens; 5, style ≈ stamens; 7, style > stamens. **C**. Number of style splits. **D**. Depth of style splitting: 3, deep; 5, medium; 7, shallow. **E**. Bend of style: 0, not bent; 5, partly bent; 9 clearly bent.

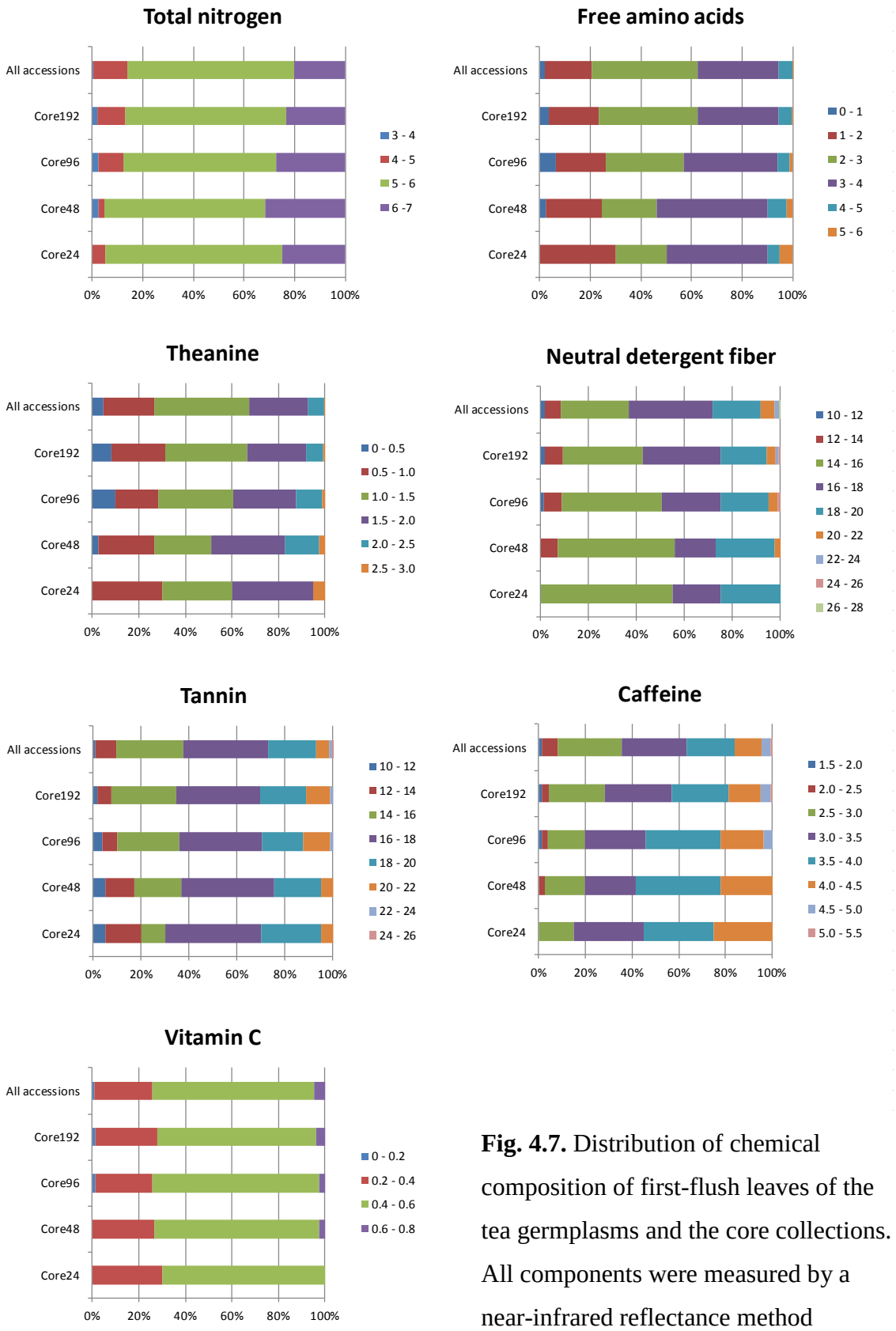

Fig. 4.7. Distribution of chemical composition of first-flush leaves of the tea germplasms and the core collections. All components were measured by a near-infrared reflectance method

CHAPTER 5 General discussion and conclusions

The objective of this study was to gather genomic information and genetic resources to improve tea breeding by enlarging the genetic diversity of breeding populations and accelerating the accumulation of useful genes.

I developed several hundred DNA markers and constructed a reference linkage map of *C. sinensis* based on them. Using the markers, I analyzed the genetic diversity of the worldwide tea germplasm collection of NIVTS, and using the genotypic data of the DNA marker loci I selected potential core collections. The core collections will support the efficient screening of new breeding materials, and the collected genomic information will allow the identification of important agronomic traits and support DNA-marker-assisted selection.

5.1 Genomic information for tea developed in this study

DNA markers used for the analysis of genetic diversity and for genetic analysis must be highly polymorphic, transferrable to various cultivars, and numerous enough to cover the entire genome. I focused on SSR markers. As SSR markers are multiallelic, it is easy to obtain markers specific to a particular allele. In addition, most of SSR markers are co-dominant and their detection is stable. Before this study, only 130 EST-SSR markers (Jin et al. 2006, Sharma et al. 2011, Zhou et al. 2011) and 49 genomic SSRs had been reported (Chen et al. 2000, Freeman et al. 2004, Hung et al. 2008, Yang et al. 2009). I developed more than 250 EST-SSR markers and 500 genomic SSR markers.

Linkage maps, which are a basic resource for genetic research, must cover the entire genome, have enough markers, and present the order of DNA markers accurately. In an F_1 population derived from cross between a Japanese green tea cultivar, ‘Sayamakaori’, and a line from China, Kana-Ck17, 441 of these markers could be mapped. The maps of the parents and the core map consisted of 15 linkage groups, the same as the basic chromosome number. Since the core map was constructed by means of integration with double the population size used in normal pseudo-testcross theory, the order of markers is more accurate than that of the parental maps. Furthermore, the markers are located evenly across the entire genome. Therefore, this map will play an

important role as a starting point for the genetic analysis of tea.

Since the average distance between DNA markers is 4.35 cM, the marker density of the reference map is not enough for fine mapping, which is necessary for map-based cloning. Furthermore, when more traits become targets of genetic analysis, it will be necessary to prepare DNA markers that are polymorphic between closely related cultivars. These approaches could need more DNA markers, so it will become important to increase the number of markers in the future. Because the genome of *C. sinensis* has not yet been sequenced, the development of novel DNA markers is laborious. However, recent progress in genome analysis made possible by the use of next-generation sequencers, such as RNA-seq analysis, will enable us to develop novel EST-SSRs or EST-SNPs more efficiently. In addition, new techniques for organisms without reference genome sequences, such as restriction-site-associated DNA sequence (Baird et al. 2008) and genotyping by sequencing (Elshire et al. 2011), will facilitate the development of more DNA markers on the linkage map of tea.

Although DNA-marker-assisted breeding is common in major annual crops, despite many genome analyses and the development of DNA markers linked to important agronomic traits in tree crops, only one tea cultivar has been developed so far by this method: ‘Nanmei’, a green tea cultivar developed by using a DNA marker for the mulberry scale resistance gene *MSR1* (Taniguchi et al. 2013). Selection markers for various traits are needed in tea breeding. So far, phenotyping methods and segregating populations have been developed for resistance to anthracnose (Yoshida et al. 2006) and for caffeine-less tea (Ogino et al. 2009). The genome resource developed in this study will be useful for these and other traits.

5.2 Worldwide core collection of tea genetic resources

Today, access to genetic resources is becoming restricted because of the growing awareness of the need to protect genetic resources. This makes the collection of further genetic resources difficult. Such restrictions make the core collections selected in this study a precious resource for tea breeding.

Takeda (2002) comprehensively phenotyped tea germplasms in NIVTS, and showed that the collection has extensive diversity in various traits, with differences in

variation in several traits among varieties or populations. However, the genetic background and relationships of each population were not clear, and the traits analyzed were those for which large-scale analysis is easy. In contrast, information about traits for which phenotyping is laborious, such as pest resistance or tea quality, was poor.

The core collections selected in this study now enable breeders to analyze variation in these traits and to screen breeding materials. The value of core collections is determined by whether breeders can easily and efficiently identify breeding materials with a desired phenotype. In our analysis of the chemical components of tender shoots, the differences in the CVs and the range of contents of each component indicated that the core collections supply enough diversity in a limited number of accessions. Other traits should also be amenable to such screening.

The core collections are linked to genetic information, such as the results of Structure analysis, PCOA, and phylogenetic analysis. If screening of the core collections revealed several candidate accessions, breeders could select the most suitable one by estimating their origins and relationships from this information.

One disadvantage of core collections is that they cannot capture rare alleles. In some cases, this might mean that breeders cannot find materials with a rare phenotype. However, the DNA marker for the mulberry scale resistance gene *MSR1*, which was detected in ‘Sayamakaori’ (var. *sinensis*), was also detected in ‘Kanaya 13’, which is unrelated to ‘Sayamakaori’, and of which the resistance gene was derived from var. *assamica* (Tanaka et al. 2004). The main core collection, core 192 showed very high coverage of genetic diversity using the data of SSR marker loci. Therefore, even a seemingly rare allele can be detected by using the core collections, a widely representative subset of genetic resources.

In this study, genetic analysis using many SSR loci revealed a new detail of the genetic diversity of tea. Matsumoto et al. (2002) concluded from polymorphisms of a CAPS marker of the phenylalanine ammonia-lyase gene that the genetic diversity of Japanese landraces, in which only three alleles were detected, is much lower than that of Chinese germplasms. However, among the Japanese accessions that I analyzed, H_E = 0.70, H_O = 0.55, and allelic richness per three accessions was 3.29 (Table 4.1), and thus they have some genetic diversity. Therefore, even though few tea seeds were brought from China to Japan, their introduction several times from different sources formed the

basis for the genetic diversity. In the context of germplasms with other origins, this study includes the first large-scale analysis of genetic diversity in a worldwide germplasm collection that used genotypic data. The findings will be important to the collection and management of tea genetic resources.

5.3 Future prospects for tea breeding

The genomic resources and core collections assembled in this study will significantly advance tea breeding. DNA-marker-assisted selection, which requires the genetic analysis of target traits and versatile linkage maps, is of great help in selecting useful genes efficiently. After the identification of target loci, core collections allow the analysis of genetic diversity. For example, in the case of resistance genes, information about gene frequency or allele diversity can assist in screening of alleles or in analysis of the relationship between resistance alleles and disease races.

Although the resources developed in this study will accelerate the identification of useful breeding materials and loci for important traits, most of the germplasms are inferior in many traits, including tea quality, yield, and cold resistance. Since *C. sinensis* is self-incompatible and shows strong inbreeding depression, normal backcrossing to a recurrent parent is difficult. This makes it necessary to perform pseudo-backcrossing to different recurrent parents in each generation. To do this efficiently, the generation time needs to be shortened from the usual minimum of 4 years. Flowering of tea can be induced by root restriction or treatment with dwarfing agents. In this way, 2-year-old plants can be made to flower. If 1-year-old plants can be made to flower, the generation time can be shortened to 2 years, facilitating pseudo-backcross breeding. Thus, the development of new flowering promotion techniques is an important issue in the future.

In pseudo-backcross breeding, selection by phenotype takes a long time and is not efficient. Therefore, selection by genotype is needed. Genomic selection, which selects superior individuals by using a breeding value estimated from the genome-wide genotype, is suitable for the selection of traits such as tea quality and yield. In several tree crops, such as Japanese pear and forest trees (Iwata et al. 2013, Beaulieu et al. 2014), research into the application of genomic selection to breeding has begun. To apply genomic selection in tea, the extent of linkage disequilibrium in core collections

or elite cultivars must first be clarified.

5.4 Conclusions

This study involved the development of DNA markers and linkage maps of tea and the streamlining of genetic resources to select core collections. From EST analysis of seven plant organs, SSR markers were developed. Genomic SSR markers were also been developed from an SSR-enriched genomic library. A reference linkage map of *C. sinensis* based on the SSR markers was constructed. The SSR markers were used to analyze the genetic diversity of the worldwide germplasm collection at NIVTS, and core collections were selected from the full collection. The core collections will facilitate the identification of useful germplasms, resulting in the improvement of genetic diversity of breeding populations of tea. The genomic information will enable breeders to capture and select agronomically important genes.

This study provides a technical basis for continuous progress in tea breeding. It will contribute to the development of new cultivars to supplant ‘Yabukita’ and revolutionize the tea industry in Japan.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my supervisor, Associate Professor Dr. Junichi Tanaka, for his kind guidance and support throughout my research and writing.

I wish to thank the committee members—Professors Ryoji Takahashi, Chikako Otobe, and Ryo Ohsawa—for their encouragement, insightful comments, and challenging questions. I also thank Professor Toshiya Yamamoto for his helpful and critical comments.

It is a pleasure to thank Dr. Hiroyuki Fukuoka for his collaboration on the work of Chapter 2 and 3 and his important advice on genome research of tea. I would like to thank Dr. Kazumi Furukawa, Ms. Sakura Ota-Metoku, Dr. Nobuo Yamaguchi, Dr. Tomomi Ujihara and Dr. Izumi Kono for their collaboration on the work of the linkage-map construction. I would also like to thank Dr. Satoshi Yamaguchi and Ms. Keiko Kimura who supported the work of Chapter 4.

I am indebted to my many colleagues who supported me at the Makurazaki Tea Research Station, NARO Institute of Vegetable and Tea Science. In particular, Mr. Atsushi Nesumi and Dr. Yoshiyuki Takeda gave me clear and important guidance and conveyed the importance of tea breeding. I would also like to thank Dr. Katsuyuki Yoshida, Dr. Akiko Ogino, Dr. Hiroshi Yorozuya and Mr. Tetsuji Saba for their collaboration on tea breeding in NIVTS. I acknowledge the technical assistance of Ms. Yuko Kuramae, Ms. Miwa Iwata and Ms. Naoko Rokutan. All the other members of the Makurazaki Tea Research Station have helped my research and the writing of this thesis.

Last but not least, I would like to thank my family for the support they have provided me through my life so far. Especially, I would like to mention about my wife, Sachiko and my son, Hironori. I greatly thank Sachiko for her understanding, patience and support. I thank Hironori for making me so happy with his cute smile and encouraging me lot at home.

REFERENCES

Anan, T. and M. Nakagawa (1974) Effect of light on chemical constituents in the tea leaves. Nippon Nogeikagaku Kaishi 48: 91–98 (in Japanese English summary).

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill L. Issel-Tarver, A. Kasarskis, S. Lewis, J. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin and G. Sherlock. (2000) Gene Ontology: tool for the unification of biology. Nat. Genet. 25: 25–29.

Ashihara, H., H. Sano and A. Crozier (2008) Caffeine and related purine alkaloids: biosynthesis, catabolism, function and genetic engineering. Phytochemistry 69: 841–856.

Baird, N.A., P.D. Etter, T.S. Atwood, M.C. Currey, A.L. Shiver, Z.A. Lewis, E.U. Selker and E.A. Johnson (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3(10): e3376.

Balfourier, F., V. Roussel, P. Strelchenko, F. Exbrayat-Vinson, P. Sourdille, G. Boutet, J. Koenig, C. Ravel, O. Mitrofanova, M. Beckert and G. Charmet (2007) A worldwide bread wheat core collection arrayed in a 384-well plate. Theor. Appl. Genet. 114: 1265–1275.

Beaulieu, J., T. Doerksen, S. Clément, J. Mackay and J. Bousquet (2014) Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. Heredity, 113: 343–352.

Campana, M.G., H.V. Hunt, H. Jones and J. White (2011) CorrSieve: software for summarizing and evaluating Structure output. Mol. Ecol. Resour. 11: 349–352.

Chakraborty, R. and L. Jin (1993) Determination of relatedness between individuals using DNA fingerprinting. Hum. Biol. 65: 875–895.

Chen, B. (2004) The origination of Tea-Horse Road. Thinking 4: 10.

Chen, L., Z.X. Zhou and Y.J. Yang (2007) Genetic improvement and breeding of tea plant (*Camellia sinensis*) in China: from individual selection to hybridization and molecular breeding. Euphytica, 154: 239-248

Chen, J., P.S. Wang, Y.M. Xia, M. Xu and S.J. Pei (2005) Genetic diversity and differentiation of *Camellia sinensis* L. (cultivated tea) and its wild relatives in

Yunnan province of China, revealed by morphology, biochemistry and allozyme studies. Genet. Resour. Crop. Evol.. 52: 41–52

Chen, L., L.P. Zhao and Q.K. Gao (2005) Generation and analysis of expressed sequence tags from the tender shoots cDNA library of tea plant (*Camellia sinensis*). Plant Sci. 168: 359–363.

Chen, Z. and L. Chen (2012) Delicious and healthy tea: an overview. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 1–11.

Chen, Z.Y., Y.S. Jiang, Z.F. Wang, J.Q. Wei, X. Wei, H. Tang and Z.C. Li (2000) Development and characterization of microsatellite markers for *Camellia nitidissima*. Conserv. Genet. 11: 1163–1165.

Cockerham, C.C. and B.S. Weir (1993) Estimation of gene flow from F-statistics. Evolution 47: 855–863.

Conesa, A., S. Götz, J. García-Gómez, J. Terol, M. Talón and M. Robles (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

Cooper, H.D., C. Spillane and T. Hodgkin (2001) Broadening the genetic base of crops: an overview. In: Cooper, H.D., C. Spillane and T. Hodgkin (eds) Broadening the Genetic Base of Crop Production. CABI, Oxfordshire, UK, pp 1–23.

Das, S.C., S. Das and M. Hazarika (2012) Breeding of tea plant (*Camellia sinensis*) in India. In: Chen L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 69–124.

Díez, C.M., A. Imperato, L. Rallo, D. Barranco and I. Trujillo (2012) Worldwide core collection of olive cultivars based on simple sequence repeat and morphological markers. Crop Sci. 52: 211–221.

Don, R.H., B. J. Cox, B. J. Wainwright and J. S. Mattick (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19: 4008.

Dorigo, M., V. Maniezzo and A. Colorni (1996) The ant system: optimization by a colony of cooperation agents. IEEE Trans. Syst. Man. Cybern. B26: 29–41.

Ellis, J.R. and J.M. Burke (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99: 125–132.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and

S.E. Mitchell (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6: e19379.

Ercisli, S. (2012) The tea industry and improvement in Turkey. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 309–321.

Evanno, G., S. Regnaut and J. Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2640.

Excoffier, L., P.E. Smouse and J.M. Quattro (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Fang, W., H. Cheng, Y. Duan, X. Jiang and X. Li (2012) Genetic diversity and relationship of clonal tea (*Camellia sinensis*) cultivars in China as revealed by SSR markers. Plant Syst. Evol. 298: 469–483.

Forrest, G.I. and D.S. Bendall (1969) The distribution of polyphenols in the tea plant (*Camellia sinensis* L.). Biochem. J. 113: 741–755.

Frankel, O.H.(1984) Genetic perspective of germplasm conservation. In: Arber, W., K. Limensee, W.J. Peacock and P. Stralinger (eds) Genetic manipulations: Impact on man and society, Cambridge University Press, Cambridge, pp161-470.

Freeman, S., J. West, C. James, V. Lea and S. Mayes (2004) Isolation and characterization of highly polymorphic microsatellites in tea (*Camellia sinensis*). Mol. Ecol. Notes 4: 324–326.

Friedman, M. (2007) Overview of antibacterial, antitoxin, antiviral, and antifungal activities of tea flavonoids and teas. Mol. Nutr. Food Res. 51: 116–134.

Fuchinoue, Y., M. Tanaka, S. Funakoshi and N. Okano (1972) A newly registered variety ‘Sayamakaori’ for green tea. Tea Res. J. 37: 12–23 (in Japanese with English summary).

Fukuoka, H., T. Nunome, Y. Minamiyama, I. Kono, N. Namiki and A. Kojima (2005) Read2Marker: a data processing tool for microsatellite marker development from a large data set. Biotechniques 39: 472–476.

Gish, W. and D.J. States (1993) Identification of protein coding regions by database similarity search. Nature Genet. 3: 266–272.

Goudet, J. (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J. Hered. 85: 485–486.

Grattapaglia, D. and R. Sederoff (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. Genetics 137: 1121–1137.

Gunasekare, M.T.K. (2012) Tea plant (*Camellia sinensis*) breeding in Sri Lanka. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 125–176.

Hackett C.A., F.N. Wachira, S. Paul, W. Powell and R. Waugh (2000) Construction of a genetic linkage map for tea (*Camellia sinensis*). Heredity 85: 346–355.

Hashimoto, M. (2001) The origin of the tea plant. In: Proc. 2001 Int. Conf. on O-Cha (Tea) Culture and Science (Session II), Shizuoka, Japan, pp J5–J7.

Hashimoto, M. and S. Takashi (1978) Morphological studies on the origin of the tea plant. V: A proposal of one place of origin by cluster analysis. Jpn. J. Trop. Agric. 21: 93–101.

Higdon, J.V. and B. Frei (2003) Tea catechins and polyphenols: health effects, metabolism, and antioxidant functions. Crit. Rev. Food Sci. Nutr. 43: 89–143.

Hung, C.Y., K.H. Wang, C.C. Huang, X. Gong, X.J. Ge and T. Y. Chiang (2008). Isolation and characterization of 11 microsatellite loci from *Camellia sinensis* in Taiwan using PCR-based isolation of microsatellite arrays (PIMA). Conserv. Genet. 9: 779-781.

Inazuka, M., T. Tahira and K. Hayashi (1996) One-tube post-PCR fluorescent labeling of DNA fragments. Genome Res. 6: 551–557.

Iwata, H., T. Hayashi, S. Terakami, N. Takada, Y. Sawamura and T. Yamamoto (2013) Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. Breed. Sci. 63: 125‑140.

Iwata, H. and S. Ninomiya (2006) AntMap: constructing genetic linkage maps using an ant colony optimization algorithm. Breed. Sci. 56: 371–377.

Jacobs, E.M. (2009) In pursuit of pepper and tea: the story of the Dutch East India Company. Netherlands Maritime Museum.

Jakobsson, M. and N.A. Rosenberg (2007) CLUMPP: a cluster matching and

permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Jeong, B.C. and Y.G. Park (2012) Tea plant (*Camellia sinenis*) breeding in Korea. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 263–288.

Jin, J.Q., H.R. Cui, W.Y. Chen, M.Z. Lu, Y.L. Yao, Y. Xin and X.C. Gong (2006) Data mining for SSRs in ESTs and development of EST-SSR marker in tea plant (*Camellia sinensis*). J. Tea Sci. 26: 17–23.

Juneja, L. R., Chu, D. C., Okubo, T., Nagato, Y. and H. Yokogoshi (1999) L-Theanine—a unique amino acid of green tea and its relaxation effect in humans. Trends Food Sci. Technol. 10: 199–204.

Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., Katayose, Y., Harada, K., Duncan, A.V. and N. Tomooka (2012) Evaluation of soybean germplasm conserved in NIAS Genebank and development of mini core collections. Breed. Sci. 61: 566–592.

Khan A.Q. (2012) Tea germlasm and inprovement in Bangladesh. In: Chen L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 289-297.

Kalia, R.K., Rai, M.K., Kalia, S., Singh, R. and A.K. Dhawan (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177: 309–334.

Kamunya, S.M., F.N. Wachira, R.S. Pathak, R.C. Muoki and R.K. Sharma (2012) Tea Improvement in Kenya. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 177–226.

Karasuyama, M. and J. Matsumoto (1988) Effect of shade treatment on the utilization of absorbed nitrogen in tea plants. Jpn. J. Soil Sci. 59: 486–492 (in Japanese with English summary).

Kaundun, S.S. and S. Matsumoto (2002) Heterologous nuclear and chloroplast microsatellite amplification and variation in tea, *Camellia sinensis*. Genome 45: 1041–1048.

Kaundun, S.S. and S. Matsumoto (2003) Development of CAPS markers based on three key genes of the phenylpropanoid pathway in tea, *Camellia sinensis* (L.) O. Kuntze, and differentiation between *assamica* and *sinensis* varieties. Theor. Appl.

Genet. 106: 375–383.

Kaundun, S.S. and Y.G. Park (2002) Genetic structure of six Korean tea plantations as revealed by RAPD-PCR markers. Crop Sci. 42: 594–601.

Kojima, Y., Ebana, K., Fukuoka, S., Nagamine, T. and M. Kawase (2005) Development of an RFLP-based rice diversity research set of germplasm. Breed. Sci. 55: 431–440.

Kukita, Y. and K. Hayashi (2002) Multicolor post-PCR labeling of DNA fragments with fluorescent ddNTPs. BioTechniques 33: 502–506.

Kuroda, Y. and Y. Hara (1999) Antimutagenic and anticarcinogenic activity of tea polyphenols. Mutat. Res.-Rev. Mutat. 436: 69–97.

Lai, J.A., W.C. Yang and J.Y. Hsiao (2001) An assessment of genetic relationship in cultivated tea clones and native wild tea in Taiwan using RAPD and ISSR markers. Bot. Bull. Acad. Sin. 42: 93–100.

Langella, O. (1999) Populations 1.2.31: a population genetic software. CNRS UPR9034. http://bioinformatics.org/populations/ Accessed 8 January 2014.

Le Cunff, L., Fournier-Level, A., Laucou, V., Vezzulli, S., Lacombe, T., Adam-Blondon, A.F., Boursiquot, J.M. and P. This (2008) Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. *sativa*. BMC Plant Biol. 8: 31.

Li, H.L. (1983) The domestication of plants in China: ecogeographical considerations. In Keightley DN (ed.) The Origins of Chinese Civilization. University of California Press, Oakland, pp 21–63.

Liu, K. and S.V. Muse (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128–2129.

Liu, Y.G., N. Misukawa, T. Oosumi and R.F. Whittier (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8: 457–463.

Maeda-Yamamoto, M., M. Sano, N. Matsuda, T. Miyase, K. Kawamoto, N. Suzuki, M. Yoshimura, H. Tachibana and K. Hakamata (2001) The change of epigallocatechin-3-*O*-(3-*O*-methyl) gallate content in tea of different varieties, tea seasons of crop and processing method. J. Jpn. Soc. Food Sci. Technol. 48: 64–68.

Matsumoto, S., Y. Kiriiwa and Y. Takeda (2002) Differentiation of Japanese green tea

cultivars as revealed by RFLP analysis of phenylalanine ammonia-lyase DNA. Theor. Appl. Genet. 104: 998–1002.

Morita, A., H. Horie, Y. Fujii, N. Takatsu, N. Watanabe, A. Yagi and H. Yokota (2004) Chemical forms of aluminum in xylem sap of tea plants (*Camellia sinensis* L.). Phytochemistry 65: 2775–2780.

Morita, A., O. Yanagisawa, S. Takatsu, S. Maeda and S. Hiradate (2008) Mechanism for the detoxification of aluminum in roots of tea plant (*Camellia sinensis* (L.) Kuntze). Phytochemistry 69: 147–153.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Ngoc D.V. (2012) Breeding of tea plant (*Camellia sinensis*) in Vietnam. In: Chen L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 241–262.

Ogino, A., J. Tanaka, K. Yoshida, F. Taniguchi, H. Omae, A. Nesumi, T. Saba, T. Takyu and Y. Takeda (2005) New parental line 'Cha Chuukanbohon Nou 6' for anthocyanin-rich tea. Bull. Natl. Inst. Veg. Tea Sci. 4: 77–85 (in Japanese).

Ogino, A., J. Tanaka, F. Taniguchi, M.R. Yamamoto and K. Yamada (2009) Detection and characterization of caffeine-less tea plants originated from interspecific hybridization. Breed. Sci. 59: 277–283.

Ohsako, T., T. Ohgushi, H. Motosugi and K. Oka (2008) Microsatellite variability within and among local landrace populations of tea, *Camellia sinensis* (L.) O. Kuntze, in Kyoto, Japan. Genet. Resour. Crop Evol. 55: 1047–1053.

Omura M., T. Ueda, M. Kita, A. Komatsu, Y. Takanokura, T. Shimada, T. E. Inagaki, H. Nesumi and T. Yoshida (2000) EST mapping of citrus. Proc. Int. Soc. Citriculture IX Congress 1: 71–74.

Ota, S. and J. Tanaka (1999) RAPD-based linkage mapping using F_1 segregating populations derived from crossings between tea cultivar 'Sayamakaori' and strain Kana-Ck17. Breed. Res. 1 (Sup. 1): 16 (in Japanese).

Park, J., J. Kim, B. Hahn, K. Kim, S. Ha and Y. Kim (2004) EST analysis of genes involved in secondary metabolism in *Camellia sinensis* (tea), using suppression subtractive hybridization. Plant Sci. 166: 953–961.

Peakall, R. and P. Smouse (2012) GenAlEx 6.5: Genetic Analysis in Excel. Population

genetic software for teaching and research—an update. Bioinformatics 82: 2537–2539.

Pflieger, S., V. Lefebvre and M. Causse (2001). The candidate gene approach in plant genetics: a review. Molecular Breeding, 7: 275-291.

Pritchard, J.K., M. Stephens and P. Donnelly (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Raina, S.N., P.S. Ahuja, R.K. Sharma, S.C. Das, P. Hardwaj, R. Negi, V. Sharma, S.S. Singh, R.K. Sud, R.K. Kalia, et al. (2012) Genetic structure and diversity of India hybrid tea. Genet. Resour. Crop Evol. 59: 1527–1541.

Romualdi, C., S. Bortoluzzi, F. D'Alessi and G.A. Danieli (2003) IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. Physiol. Genomics 12: 159–162.

Rosenberg, N.A. (2004) Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4: 137–138.

Saijo, R. (1980) Effect of shade treatment on biosynthesis of catechins in tea plants. Plant Cell Physiol. 21: 989–998.

Sano, M., M. Suzuki, T. Miyase, K. Yoshino and M. Maeda-Yamamoto (1999) Novel antiallergic catechin derivatives isolated from oolong tea. J. Agric. Food Chem. 47: 1906–1910.

Sharma, P and S. Kumar (2005) Differential display-mediated identification of three drought-responsive expressed sequence tags in tea [*Camellia sinensis* (L.) O. Kuntze]. J. Biosci. 30: 231–235.

Sharma, H., R. Kumar, V. Sharma, V. Kumar and P. Bhardwaj (2011) Identification and cross-species transferability of 112 novel unigene-derived microsatellite markers in tea (*Camellia sinensis*). Am. J. Bot. 98: e133–138.

Sharma, R., P. Bhardwaj, R. Negi, T. Mohapatra and P. Ahuja (2009) Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). BMC Plant Biol. 9: 53.

Shi, C.Y., H. Yang, C.L. Wei, O. Yu, Z.Z. Zhang, C.J. Jiang, J. Sun, Y.Y. Li, Q. Chen, T. Xia and X.C. Wan (2011) Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. BMC Genomics 12: 131.

Shi, T., H. Huang and M.S. Barker (2010) Ancient genome duplications during the evolution of kiwifruit (*Actinidia*) and related Ericales. Ann. Bot. 106: 497–504.

Sriyadi, B., R. Suprihatini and H.S. Khomaeni (2012) The development of high yielding tea clones to increase Indonesian tea production. In: Chen L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 299–308.

Takeda, Y. (2002) Studies on variations in genetic resources of tea in Japan and application to tea breeding. Bull. Natl. Inst. Veg. Tea Sci. 1: 97–180 (in Japanese with English summary).

Takeda, Y. and T. Toyao (1980) Identification and classification of the green tea varieties by the morphological characters of flower organs. Tea Res. J. 52: 1–6 (in Japanese with English summary).

Takeo, T. (1992) Green and semi-fermented teas. In: Willson, K.C. and M.N. Clifford (eds) Tea: Cultivation to Consumption. Chapman & Hall, London, pp 413–457.

Tanaka, J. (2006) Breeding system combining of juvenile selection using DNA markers and acceleration of generations, aiming for the significant improvement of the tea breeding in Makurazaki Tea Research Station, NIVTS. Breed. Res. 8: 119–126.

Tanaka, J. (2012) Japanese tea breeding history and the future perspective. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 227–239.

Tanaka, J. and S. Ikeda (2000) Generation of microsatellite markers by PCR in tea. Breed. Res. 2 (Sup. 1): 37 (in Japanese).

Tanaka, J. and S. Ikeda (2002) Rapid and efficient DNA extraction method from various plant species using diatomaceous earth and a spin filter. Breed. Sci. 52: 151–155.

Tanaka, J. and Y. Sawai (2005) *Mat* locus control the existence of an aromatic substance methyl anthranilate in tea leaves, derived from clonal strain ‘Shizu-Inzatsu 131’. In Res. Results of Veg. & Tea Sci., NARO, Tsukuba, Japan, 75–76 (In Japanese).

Tanaka, J and F. Taniguchi (2007) Tea. In: Kole C (ed.) Genome Mapping and Molecular Breeding in Plants, vol. 6, Technical Crops. Springer-Verlag, Berlin, pp 119–125.

Tanaka, J., Y. Sawai and S. Yamaguchi (1995) Genetic analysis of RAPD markers in

tea. Jpn. J. Breed. 45 (Sup. 2): 198 (in Japanese).

Tanaka, J., N. Yamaguchi and Y. Nakamura (2001) Pollen parent of tea cultivar Sayamakaori with insect and cold resistance may not exist. Breed. Res. 3: 43–48 (in Japanese with English summary).

Tanaka, J., S. Ota-Metoku and Y. Takeda (2003) The garden-variety camellia 'Robiraki' derived from crossing between *Camellia japonica* as a seed parent and *C. sinensis* as a pollen parent. Application of RAPD and an SSR marker analysis to tea breeding by interspecific hybridization. Breed. Res. 5: 149–154 (in Japanese with English summary).

Tanaka, J., M. Takeda, M. Takashi, F. Taniguchi, M. Shige, Y. Uezono and S. Ikeda (2004) Marker-assisted selection (MAS) in tea for the mulberry scale resistance using rapid DNA extraction and emphasized-RAPD (e-RAPD) system. In: Proc. 2004 Int. Conf. on O-Cha (Tea) Culture and Science, Shizuoka, Japan, pp 149–152.

Tanaka, J., F. Taniguchi, N. Hirai and S. Yamaguchi (2006) Estimation of the genome size of tea (*Camellia sinensis*), camellia (*C. japonica*), and their interspecific hybrids by flow cytometry. Tea Res. J. 101: 1–7.

Taniguchi, F (2013) Recent advances in tea breeding and development of Japanese tea products from new cultivars. In: Proc. 2013 Int. Conf. on O-Cha (Tea) Culture and Science, Shizuoka, Japan, PR-S-4.

Taniguchi, F., K. Yoshida, J. Tanaka and A. Ogino (2010) Identification of quantitative trait loci for resistance to anthracnose in the tea cultivar 'Minamisayaka' using an SSR-based linkage map. In: Proc. 2010 Int. Conf. on O-Cha (Tea) Culture and Science, Shizuoka, Japan, PR-P-04.

Taniguchi, F., K. Yoshida, H. Yorozuya, A. Ogino, T. Saba and A. Nesumi (2013) Use of marker-assisted selection to develop a new tea cultivar, 'Nanmei', resistant to mulberry scale. In: Proc. 2013 Int. Conf. on O-Cha (Tea) Culture and Science, Shizuoka, Japan, PR-P-30.

Tanksley, S.D. and S.R. McCough (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063–1066.

Teramoto, M. (2001) Quantitative analysis of the recent trends of green tea consumption. Keizaigaku Ronkyu 55; 41-72.

Tomo, J., Y. Fuchinoue and H. Fuchinoue (1956) Study of self-fertilization of the tea plant. Tea Res. J. 7: 14–20 (in Japanese with English summary).

Toyao, T. and Y. Takeda (1999) Studies on geographical diversity of floral morphology of tea plant (*Camellia sinensis* (L.) O. Kuntze) using the method of numerical taxonomy. Tea Res. J. 87: 39–57 (in Japanese with English summary).

Tsuji, M., T. Kuboi and S. Konishi (1994) Stimulatory effects of aluminum on the growth of cultured roots of tea. Soil Sci. Plant Nutr. 40: 471–476.

Ujihara T., F. Taniguchi, J. Tanaka and N. Hayashi (2011) Development of expressed sequence tag (EST)-based cleaved amplified polymorphic sequence (CAPS) markers of tea plant and their application for cultivar identification. J. Agric. Food Chem. 59: 1557–1564.

Vigouroux, Y., J.C. Glaubitz, Y. Matsuoka, M.M., Goodman, G.J. Sánchez and J. Doebley (2008) Population structure and genetic diversity of new world maize races assessed by DNA microsatellites. Am. J. Bot. 95: 1240–1253.

Voorrips, R.E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. Heredity 93: 77–78.

Wachira, F.N., J. Tanaka and Y. Takeda (2001) Genetic variation and differentiation in tea germplasm revealed by RAPD and AFLP variation. J. Hort. Sci. Biotechnol. 76: 557–563.

Wang, X.C., L. Chen and Y. Yang (2011) Establishment of core collection for Chinese tea germplasm based on cultivated region grouping and phenotypic data. Front. Agric. China 5: 344–350.

Yamaguchi, N. and J. Tanaka (1999) Classification of Japanese original tea by maternal inherited RAPDs. Breed. Res. 1 (Sup. 2): 295 (in Japanese).

Yamamoto, T., T. Kimura, M. Shoda, T. Imai, T. Saito, Y. Sawamura, K. Kotobuki, T. Hayashi and N. Matsuta (2002) Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. Theor. Appl. Genet. 106: 9–18.

Yamamoto, T., M. Yamaguchi and T. Hayashi (2005) An integrated genetic linkage map of peach by SSR, STS, AFLP and RAPD. J. Jpn. Soc. Hort. Sci. 74: 204–213.

Yamamoto T., T. Kimura, S. Terakami, C. Nishitani, Y. Sawamura, T. Saito, K. Kotobuki and T. Hayashi (2007) Integrated reference genetic linkage map of pear

based on SSR and AFLP markers. Breed. Sci. 57: 321–329.

Yamanishi, T. (1995) Special issue on tea. Food Rev. Int. 11: 371–546.

Yang JB, J. Yang, H.T. Li, Y. Zhao and S.X. Yang (2009) Isolation and characterization of 15 microsatellite markers from wild tea plant (*Camellia taliensis*) using FIASCO method. Conserv. Genet. 10: 1621–1623.

Yao, M.Z. and L. Chen (2012) Tea germplasm and breeding in China. In: Chen, L., Z. Apostolides and Z.M. Chen (eds) Global Tea Breeding, Springer, Berlin, pp 13–68.

Yao, M.Z., L. Chen and Y.R. Liang (2008) Genetic diversity among tea cultivars from China, Japan and Kenya revealed by ISSR markers and its implication for parental selection in tea breeding programs. Plant Breed. 127: 166–172.

Yao, M.Z., C.L. Ma, T.T. Qiao, J.Q. Jin and L. Chen (2012) Diversity distribution and population structure of tea germplasms in China revealed by EST-SSR markers. Tree Genet. Genomes 8: 205–220.

Yoshida, K. and Y. Takeda (2006) Evaluation of anthracnose resistance among tea genetic resources by wound-inoculation assay. Jpn. Agric. Res. Quart. 40: 379–386.

Zhou, Y, X. Qiao, C. Ma, T. Qiao, J. Jin, M. Yao and L. Chen (2011) Genetic diversity and structure of tea landraces from guangxi based on EST-SSR analysis. Sci Silvae Sin 47: 59–67.

LIST OF PUBLICATIONS

Taniguchi, F., H. Fukuoka and J. Tanaka (2012) Expressed sequence tags from organ-specific cDNA libraries of tea (*Camellia sinensis*) and polymorphisms and transferability of EST-SSRs across *Camellia* species. Breed. Sci. 62(2): 186–195.

Taniguchi, F., K. Furukawa, S. Ota-Metoku, N. Yamaguchi, T. Ujihara, I. Kono, H. Fukuoka and J. Tanaka (2012) Construction of a high-density reference linkage map of tea (*Camellia sinensis*). Breed. Sci. 62(3): 263–273.

Taniguchi, F., K. Kimura, T. Saba, A. Ogino, S. Yamaguchi and J. Tanaka (2014) Worldwide core collections of tea (*Camellia sinensis*) based on SSR markers. Tree Genet Genomes, doi: 10.1007/s11295-014-0779-0.